

(19) RÉPUBLIQUE FRANÇAISE  
INSTITUT NATIONAL  
DE LA PROPRIÉTÉ INDUSTRIELLE  
PARIS

(11) N° de publication : 2 780 733

(à n'utiliser que pour les  
commandes de reproduction)

(21) N° d'enregistrement national : 99 08141

(51) Int Cl<sup>7</sup> : C 12 N 15/31, C 12 N 15/63, 5/10, C 07 K 14/195,  
A 23 K 1/00, A 23 C 19/00, A 61 K 38/43, 48/00, C 12 P 19/42

(12) DEMANDE DE BREVET D'INVENTION

A1

(22) Date de dépôt : 25.06.99.

(30) Priorité : 25.06.98 EP 98305033.

(43) Date de mise à la disposition du public de la  
demande : 07.01.00 Bulletin 00/01.

(56) Liste des documents cités dans le rapport de  
recherche préliminaire : *Ce dernier n'a pas été  
établi à la date de publication de la demande.*

(60) Références à d'autres documents nationaux  
apparentés :

(71) Demandeur(s) : GIST-BROCADES B.V. — NL.

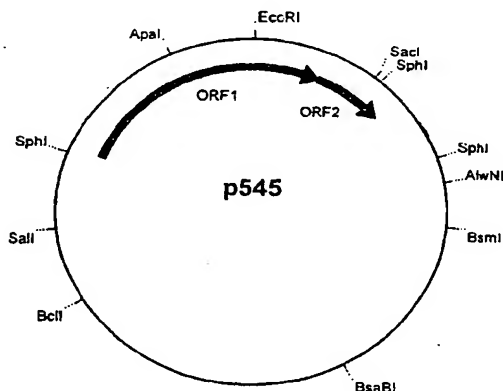
(72) Inventeur(s) : POWELS PIETER HENDRIK, VAN  
LUIJK NICOLE, JORE JOHANNES PETRUS MARIA et  
LUITEN RUDOLF GIJSBERTUS MARIE.

(73) Titulaire(s) :

(74) Mandataire(s) : CABINET BEAU DE LOMENIE.

(54) POLYNUCLEOTIDE, VECTEUR ET CELLULE HOTE LE CONTENANT, POLYPEPTIDE CORRESPONDANT,  
PROCEDE POUR LEUR PRODUCTION ET LEUR UTILISATION.

(57) L'invention concerne un polynucléotide ou plasmide  
dérivé d'un plasmide endogène de *Propionibacterium* p545,  
qui peut être utilisé pour transformer des cellules hôtes de  
*Propionibacterium* pour qu'elles produisent des polypepti-  
des ou protéines homologues ou hétérologues ou d'autres  
composés, par exemple la vitamine B12, un procédé de pro-  
duction correspondant et une cellule hôte utilisée dans ce  
procédé.



FR 2 780 733 - A1



La présente invention concerne un polynucléotide comprenant une séquence spécifique, un vecteur le contenant, un polypeptide codé par ce polynucléotide, une cellule hôte comprenant ce polynucléotide ou ce vecteur, un procédé de production de cette cellule hôte, un procédé de préparation de ce polypeptide, un procédé de préparation de la vitamine B12 par culture de la cellule hôte, et l'utilisation du polypeptide dans une méthode de traitement du corps humain ou animal, et l'utilisation de la cellule hôte pour la fabrication du fromage, d'aliments ou d'aliments pour animaux.

Plus précisément, la présente invention concerne un plasmide endogène des bactéries *Propionibacterium*, des vecteurs qui en sont dérivés et l'utilisation de ces vecteurs pour exprimer des protéines (hétérologues) dans des bactéries, en particulier dans des bactéries *Propionibacterium*.

Les bactéries *Propionibacterium* sont des bactéries Gram-positives capables de produire divers composés intéressants dans différents procédés industriels. Par exemple, on sait que plusieurs espèces de *Propionibacterium* produisent la vitamine B12 (cobalamine) dans des procédés de fermentation à grande échelle. D'autres espèces sont utilisées dans les applications laitières comme la fabrication du fromage où elles contribuent à, et dans bien des cas sont même principalement responsables, de la flaveur et de la texture spécifiques du fromage. On considère que de nombreuses espèces de *Propionibacterium* peuvent être incorporées en toute sécurité, sous forme d'organismes vivants, dans les aliments et les aliments pour animaux.

Pour pouvoir exploiter totalement le potentiel biotechnologique des bactéries *Propionibacterium*, il est nécessaire de disposer de techniques de génie génétique efficaces et flexibles. De telles techniques sont basées sur la disponibilité d'un plasmide approprié pour exprimer une protéine à partir d'un gène hétérologue dans des bactéries *Propionibacterium*.

Le document EP-A-0400931 (Nippon Oil) concerne un plasmide endogène (pTY-1) provenant de *Propionibacterium pentosaceum* (ATCC 4875) mais ne décrit pas sa séquence et n'indique pas de manière concrète comment il peut être utilisé pour exprimer un gène hétérologue.

Le document JP 08-56673 concerne le plasmide pTY-1 pour la production de la vitamine B12 mais n'indique pas que le plasmide demeure un élément extrachromosomique qui se réplique librement ni que le plasmide est stable dans les cellules transformées.

C'est pourquoi la présente invention a pour but de fournir des vecteurs qui sont plus efficaces que les vecteurs de l'état de la technique, qui peuvent demeurer extrachromosomiques et qui sont stables. En particulier, la présente invention a pour but de fournir un vecteur efficace pour le clonage et l'expression  
5 de fragments génomiques étrangers ou de gènes étrangers dans une souche hôte de *Propionibacterium*. Ceci peut leur permettre de se soustraire à l'action d'enzymes de restriction spécifiques et d'éviter ainsi que l'hôte traite le plasmide comme un polynucléotide étranger.

Ainsi, dans un premier aspect, la présente invention fournit un  
10 polynucléotide comprenant une séquence capable de s'hybrider sélectivement avec  
(a) la séquence SEQ ID n° 1 ou sa séquence complémentaire ;  
(b) une séquence issue du plasmide de 3,6 kb de *Propionibacterium freudenreichii* CBS 101022 ;  
(c) une séquence issue du plasmide de 3,6 kb de *Propionibacterium freudenreichii* CBS 101023, ou  
15 (d) une séquence qui code un polypeptide de l'invention, comme (au moins une partie de) la séquence d'acides aminés représentée dans SEQ ID n° 2 ou SEQ ID n° 3 ou sa séquence complémentaire. SEQ ID n° 1 présente la séquence d'ADN du plasmide endogène de  
20 *Propionibacterium* LMG 16545 que la demanderesse a découvert. La première séquence codante s'étend du nucléotide 273 au nucléotide 1184 et la séquence d'acides aminés déduite de cette séquence codante est montrée dans SEQ ID n° 2. La seconde séquence codante s'étend du nucléotide 1181 au nucléotide 1483 et la séquence d'acides aminés  
25 déduite de cette séquence codante est montrée dans SEQ ID n° 3.

La demanderesse a étudié un grand nombre d'isolats de *Propionibacterium* et a identifié deux souches contenant des plasmides cryptiques d'une taille de 3,6 kb. L'une de ces souches est *Propionibacterium freudenreichii* LMG 16545 qui a été déposée auprès du Centraalbureau voor Schimmelcultures  
30 (CBS), Oosterstraat 1, Postbus 273, NL-3740 AG Baarn, Pays-Bas, au nom de Gist-brocades B.V. de Wateringseweg. 1, P.O. Box 1, 2600 MA Delft, Pays-Bas, le 19 juin 1998 conformément au Traité de Budapest et qui s'est vue attribuer le numéro d'ordre CBS 101022. L'autre souche est *Propionibacterium freudenreichii* LMG 16546 qui a été déposée auprès du Centraalbureau voor Schimmelcultures le  
35 19 juin 1998 conformément au Traité de Budapest et qui s'est vue attribuer le numéro d'ordre CBS 101023.

Grâce à la caractérisation totale et à l'analyse assistée par ordinateur de la séquence nucléotidique de LMG 16545, la demanderesse a pu identifier des sites d'insertion pour des fragments d'ADN étrangers. Ces sites ont permis de construire des plasmides qui sont encore capables de se répliquer de manière autonome dans la bactérie *Propionibacterium*.

De manière surprenante, on a constaté qu'un gène de résistance à l'érythromycine provenant de l'actinomycète *Saccharopolyspora erythraea* est exprimé efficacement dans *Propionibacterium* de sorte qu'il peut être utilisé comme marqueur de sélection pour des cellules transformées.

On a également construit des vecteurs bifonctionnels qui peuvent être maintenus de manière stable et qui peuvent être sélectionnés dans les bactéries *E. coli* et *Propionibacterium*. Ceci permet d'utiliser des bactéries *E. coli* pour la construction de vecteurs, et d'exprimer de manière fonctionnelle des gènes homologues ou hétérologues dans des bactéries *Propionibacterium*. La construction de vecteurs avec *E. coli* est relativement aisée et peut être réalisée rapidement.

Le polynucléotide selon l'invention peut se répliquer de manière autonome ou extrachromosomique, par exemple dans une bactérie comme *Propionibacterium*.

Ainsi, dans un second aspect, la présente invention fournit aussi un vecteur qui comprend un tel polynucléotide.

L'invention fournit également un procédé de préparation d'un polypeptide qui comprend la culture d'une cellule hôte transformée ou transfectée avec un vecteur selon l'invention dans des conditions permettant l'expression du polypeptide.

L'invention fournit en outre un polypeptide qui comprend la séquence présentée dans SEQ ID n° 2 ou 3 ou une séquence sensiblement homologue de celle-ci, ou un fragment de l'une ou l'autre séquence, ou une protéine codée par un polynucléotide selon la présente invention.

Un polynucléotide selon l'invention peut être capable de s'hybrider sélectivement avec la séquence SEQ ID N° 1, ou avec une partie de celle-ci, ou avec la séquence complémentaire de cette séquence ou d'une partie de celle-ci. Ce polynucléotide peut être capable de s'hybrider sélectivement avec la séquence du plasmide de 3,6 kb de *P. freudenreichii* CBS 101022 ou CBS 101023, ou avec une partie de la séquence de l'un ou l'autre plasmide. Typiquement, un polynucléotide selon l'invention est une séquence contiguë de nucléotides qui est capable de

s'hybrider sélectivement avec la séquence SEQ ID n° 1 ou avec la séquence de l'un ou l'autre plasmide de 3,6 kb, ou avec une partie de l'une quelconque de ces séquences, ou avec la séquence complémentaire de ces séquences ou d'une partie de l'une quelconque de ces séquences.

5           Un polynucléotide selon l'invention et la séquence SEQ ID n° 1 ou l'un ou l'autre des plasmides de 3,6 kb, ou une séquence codant un polypeptide, ou une partie de ces séquences, peuvent s'hybrider à un niveau sensiblement supérieur au bruit de fond. Une hybridation de bruit de fond peut se produire par exemple du fait de la présence d'autres polynucléotides dans la préparation. Le niveau du  
10 signal produit par l'interaction entre un polynucléotide de l'invention et la séquence SEQ ID n° 1 ou l'un ou l'autre des plasmides de 3,6 kb, ou une partie de ces séquences, est typiquement au moins 10 fois, de préférence au moins 100 fois, plus intense que les interactions entre d'autres polynucléotides et la séquence codante de SEQ ID n° 1 ou de l'un ou l'autre des plasmides de 3,6 kb, ou une  
15 séquence codant un polypeptide, ou une partie de ces séquences. On peut mesurer l'intensité d'interaction par exemple en radiomarquant la sonde, par exemple avec <sup>32</sup>P. Typiquement, on obtient une hybridation sélective en utilisant des conditions de stringence moyenne (par exemple chlorure de sodium 0,03M et citrate de sodium 0,03M à une température d'environ 50°C) à élevée (mêmes conditions  
20 mais à 60°C).

Les polynucléotides selon l'invention peuvent présenter une homologie généralement d'au moins 70 %, de préférence d'au moins 80 ou 90 % et de préférence encore d'au moins 95 %, et de manière optimale d'au moins 98 %, avec les séquences (a) à (d) sur une région d'au moins 20, de préférence d'au moins 30,  
25 par exemple d'au moins 40, 60 ou 100 nucléotides contigus ou plus.

Il est possible d'utiliser une combinaison quelconque des degrés d'homologie et des tailles minimales mentionnés ci-dessus pour définir les polynucléotides selon l'invention, les combinaisons les plus stringentes (c'est-à-dire plus haute homologie sur de plus grandes longueurs) étant préférées. Ainsi,  
30 par exemple, un polynucléotide qui a une homologie d'au moins 80 % ou 90 % sur 25, de préférence sur 30 nucléotides, constitue un mode de réalisation de l'invention, de même qu'un polynucléotide qui a une homologie d'au moins 90 % ou 95 % sur 40 nucléotides.

Les parties évoquées ci-dessus peuvent être les séquences codantes de  
35 SED ID n° 1 ou de l'un ou l'autre des plasmides de 3,6 kb. D'autres parties de SEQ ID n° 1 que l'on préfère sont l'origine de réplication, les séquences promotrices ou

régulatrices, ou les séquences capables de provoquer ou de favoriser la réplication autonome dans une cellule hôte, telle qu'une bactérie *Propionibacterium*.

On a constaté que la partie du plasmide qui s'étend du site de restriction *SaII* au site de restriction *AlwNI* semble être la région qui est nécessaire pour la réplication du plasmide. Bien que d'autres parties du plasmide aient été  
5 déléetées, la réplication ne semble pas avoir été affectée de manière défavorable. De ce fait, selon l'invention, la séquence (b) ou (c) peut être la région délimitée par les sites de restriction *SaII* et *AlwNI* qui a une longueur d'environ 1,7 kb. A titre d'alternative, la séquence (b) ou (c) peut être remplacée par la séquence  
10 correspondant aux nucléotides 1 à 1 800, par exemple 100 à 1 700, de manière appropriée 150 à 1 500, avantageusement 200 à 1 300 et de manière optimale 250 à 1 200 de SEQ ID n° 1.

Les protéines (SEQ ID n° 2 et 3) codées par ORF1 (cadre de lecture 1) et ORF 2, respectivement, sont supposées favoriser l'une et l'autre la réplication du  
15 plasmide. Le plasmide se réplique par le processus connu de réplication en cercle rotatif dans lequel l'ADN plasmidique double brin initial est coupé par l'une ou l'autre des protéines qui favorisent la production d'une copie du brin externe en utilisant le brin interne comme matrice. La copie du cercle externe est retirée et les  
20 extrémités sont réunies. L'hôte assure ensuite la réplication d'un nouveau cercle interne en utilisant comme matrice le cercle externe produit.

Les séquences codantes selon l'invention peuvent être modifiées par des substitutions de nucléotides, par exemple par 1, 2 ou 3 à 10, 25, 50 ou 100 substitutions. A titre d'alternative ou de complément, les polynucléotides des  
25 séquences (a) à (d) peuvent être modifiés par une ou plusieurs insertions ou délétions et/ou par une extension à l'une ou l'autre des extrémités. Il est possible de réaliser des substitutions dégénérées et/ou des substitutions qui entraînent une substitution d'acides aminés conservative quand la séquence modifiée est traduite, par exemple de la manière discutée plus loin au sujet des polypeptides de  
l'invention.

30 Les polynucléotides selon l'invention peuvent être de l'ADN ou de l'ARN. Ce peut être également des polynucléotides qui incluent des nucléotides synthétiques ou modifiés. Un certain nombre de types différents de modifications qui peuvent être apportées à des polynucléotides sont connus dans la technique. Celles-ci comprennent les squelettes de méthylphosphonate et de thiophosphate,  
35 l'addition d'acridine ou de chaînes de polylysine aux extrémités 3' et/ou 5' de la molécule. Il est à noter que les polynucléotides selon l'invention peuvent être

modifiés par tout procédé disponible dans la technique. De telles modifications peuvent être apportées pour augmenter l'activité *in vivo* ou la durée de vie des polynucléotides de l'invention.

Les polynucléotides selon l'invention peuvent être utilisés comme  
5 amorce, par exemple une amorce de PCR (réaction en chaîne par polymérase), une amorce pour une autre réaction d'amplification, une amorce qui peut être marquée par exemple avec un marqueur d'identification, radioactif ou non radioactif, par des moyens conventionnels, ou bien les polynucléotides peuvent être incorporés ou clonés dans des vecteurs. Ces amorces, sondes et autres fragments auront une  
10 longueur d'au moins 15, de préférence d'au moins 20, par exemple d'au moins 25, 30 ou 40 nucléotides. Ils auront typiquement une longueur pouvant atteindre 40, 50, 60, 70, 100 ou 150 nucléotides. Les sondes et fragments peuvent avoir une longueur supérieure à 150 nucléotides, par exemple pouvant atteindre 200, 300, 500, 1 000 ou 1 500 nucléotides, ou bien une longueur de quelques nucléotides,  
15 par exemple de 5 à 10 nucléotides, de l'une quelconque des séquences (a) à (d).

Les polynucléotides tels que les polynucléotides d'ADN et les amorces selon la présente invention peuvent être produits par recombinaison, par synthèse ou par tout moyen dont dispose l'homme du métier. Ils peuvent aussi être clonés par des techniques standards. Les polynucléotides sont typiquement fournis sous  
20 forme isolée et/ou purifiée.

En général, les amorces seront produites par synthèse, avec production par étapes de la séquence d'acide nucléique voulue, nucléotide par nucléotide. On dispose largement de techniques automatisées pour réaliser une telle synthèse.

Habituellement, les polynucléotides de plus grande taille seront  
25 obtenus par recombinaison, par exemple à l'aide des techniques de clonage par PCR qui comprennent la production d'une paire d'amorces (par exemple d'environ 15 à 30 nucléotides) pour la région de SEQ ID n° 1 ou de l'un ou l'autre des plasmides de 3,6 kb que l'on souhaite cloner, la mise en contact des amorces avec de l'ADN obtenu à partir d'une bactérie *Propionibacterium*, la mise en œuvre d'une  
30 réaction en chaîne par polymérase dans des conditions qui entraînent l'amplification de la région voulue, l'isolement du fragment amplifié (par exemple par purification du mélange réactionnel sur un gel d'agarose) et la récupération de l'ADN amplifié. Les amorces peuvent être conçues pour contenir des sites de restriction appropriés afin que l'ADN amplifié puisse être cloné dans un vecteur de  
35 clonage approprié. De telles techniques peuvent être utilisées pour obtenir tout ou partie de SEQ ID n° 1 ou de l'un ou l'autre des plasmides de 3,6 kb.

Les techniques mentionnées ici sont bien connues<sup>10</sup>.

Les polynucléotides qui ne présentent pas une homologie de 100 % avec SEQ ID n° 1 ou avec l'un ou l'autre des plasmides de 3,6 kb mais qui sont conformes à l'invention peuvent être obtenus de diverses manières.

5 Les polynucléotides homologues de SEQ ID n° 1 ou de l'un ou l'autre des plasmides de 3,6 kb peuvent être obtenus par exemple par criblage de banques d'ADN génomique obtenues à partir de diverses bactéries *Propionibacterium*, comme *P. freudenreichii*, *P. jensenii*, *P. thoenii*, *P. acidipropionici*, ou d'autres souches de bactéries de la classe des actinomycètes, ou d'autres bactéries Gram-  
10 positives, ou à partir de celles qui sont riches en G:C. Tous ces organismes constituent des sources appropriées de gènes homologues ou hétérologues, de promoteurs, de stimulateurs, ou de cellules hôtes, qui peuvent être utilisés selon la présente invention. De tels homologues et leurs fragments sont généralement capables de s'hybrider sélectivement avec la séquence codante de SEQ ID n° 1 ou  
15 avec sa séquence complémentaire, ou avec l'un ou l'autre des plasmides de 3,6 kb. De telles séquences peuvent être obtenues par criblage de banques d'ADN génomique de *Propionibacterium* avec des sondes comprenant tout ou partie de la séquence codante de SEQ ID n° 1 ou de l'un ou l'autre des plasmides de 3,6 kb, dans des conditions de stringence moyenne à élevée (par exemple chlorure de  
20 sodium 0,3M et citrate de sodium 0,03M à une température d'environ 50 à 60°C).

Il est possible aussi d'obtenir des homologues par PCR dégénérée qui utilise des amorces dirigées vers des séquences conservées dans les homologues. Les séquences conservées peuvent être prédites par alignement de SEQ ID n° 1 ou de la séquence de l'un ou l'autre des plasmides de 3,6 kb avec leurs homologues.  
25 Les amorces contiendront une ou plusieurs positions dégénérées et seront utilisées dans des conditions de stringence plus basses que celles utilisées pour cloner des séquences avec des amorces à une seule séquence contre des séquences connues.

A titre d'alternative, il est possible d'obtenir de tels polynucléotides par mutagenèse dirigée de SEQ ID n° 1 ou de l'un ou l'autre des plasmides de 3,6 kb, ou de leurs homologues. Ceci peut être utile par exemple quand des changements  
30 de codons silencieux sont nécessaires pour des séquences, pour optimiser les préférences de codons pour une cellule hôte particulière dans laquelle les séquences polynucléotidiques sont exprimées. D'autres changements de séquence peuvent être souhaitables pour introduire des sites de restriction ou pour modifier  
35 les propriétés ou les fonctions des polypeptides codés par les polynucléotides.



Des procédés pour mesurer l'homologie de polynucléotides sont bien connus dans la technique. Par exemple, l'ensemble UWGCG fournit le programme BESTFIT qui peut être utilisé pour calculer l'homologie, par exemple en mode de défaut<sup>7</sup>. Pour l'homologie des acides aminés en ce qui concerne les polypeptides  
5 selon l'invention dont il sera question dans la suite, on peut employer BLAST (Basic Local Alignment Search Tool<sup>1</sup>), qui produit des alignements de séquences d'acides aminés (et de séquences de nucléotides si nécessaire) pour déterminer la similitude de séquences. Ainsi, BLAST peut être utilisé pour déterminer des appariements exacts ou pour identifier des homologues, et il est particulièrement  
10 utile pour les appariements qui ne contiennent pas de vides. La technique BLAST utilise l'algorithme basé sur l'appariement de segments à haute évaluation ("High-scoring Segment Pair" (HSP)).

Les polynucléotides double brin comprenant une séquence polynucléotidique selon l'invention et sa séquence complémentaire sont inclus  
15 dans le cadre de l'invention.

Les polynucléotides (par exemple sondes ou amorces) selon l'invention peuvent porter un marqueur d'identification. Les marqueurs appropriés comprennent des radioisotopes comme <sup>32</sup>P ou <sup>35</sup>S, les marqueurs enzymatiques ou d'autres marqueurs protéiques comme la biotine. Des techniques de détection pour  
20 ces marqueurs sont connues en soi.

Les polynucléotides marqués ou non marqués, peuvent être utilisés dans des tests basés sur les acides nucléiques pour détecter ou séquencer un autre polynucléotide selon l'invention dans un échantillon.

Les polynucléotides selon l'invention comprennent les variantes de la  
25 séquence de SEQ ID n° 1 ou de l'un ou l'autre des plasmides de 3,6 kb qui sont capables de se répliquer de manière autonome ou de demeurer extrachromosomiques dans une cellule hôte. De telles variantes peuvent être stables dans une bactérie telle qu'une bactérie *Propionibacterium*.

Généralement le polynucléotide comprendra l'origine de réplication  
30 et/ou la ou les régions codantes de SEQ ID n° 1 ou de l'un ou l'autre des plasmides de 3,6 kb, ou des homologues de ces séquences. Un polynucléotide qui est stable dans une cellule hôte, telle qu'une cellule de *Propionibacterium* ou de *E. coli*, est un polynucléotide qui persiste dans l'hôte au cours de cinq générations, de préférence au cours de quinze générations et de préférence encore au cours de  
35 trente générations. Généralement, un tel polynucléotide est hérité par les deux cellules filles à chaque génération.

Le polynucléotide peut comprendre un promoteur ou une origine de réplication (par exemple en amont de toutes les séquences codant une protéine de réplication).

Le polynucléotide selon l'invention peut être utilisé pour transformer  
5 une bactérie, telle qu'une bactérie *Propionibacterium* ou *E. coli*, par exemple par un procédé approprié<sup>11</sup>.

Le polynucléotide peut être présent dans une bactérie en un nombre de copies de 5 à 500, par exemple de 10 à 100.

Le polynucléotide peut être capable de se répliquer de manière  
10 autonome dans une bactérie qui n'est pas une bactérie *Propionibacterium*. Ce peut être une bactérie *E. coli* ou une bactérie Gram-positive ou riche en G : C, ou une bactérie de la classe des actinomycètes. Un tel polynucléotide comprendra généralement des séquences qui permettent au polynucléotide de se répliquer de  
15 manière autonome dans cette bactérie. De telles séquences peuvent être issues de plasmides qui sont capables de se répliquer dans cette bactérie. Le polynucléotide peut être un polynucléotide qui a été produit par réplication dans une bactérie *Propionibacterium*. A titre d'alternative, le polynucléotide peut avoir été produit par réplication dans une autre bactérie, telle qu'une bactérie *E. coli*. Ce polynucléotide peut être capable de se soustraire à l'action des enzymes de  
20 restriction de la cellule hôte de *Propionibacterium*.

Dans un second aspect, l'invention concerne également un vecteur comprenant un tel polynucléotide. Ce vecteur peut être capable de se répliquer dans une cellule hôte, telle qu'une bactérie, par exemple actinomycète, *Propionibacterium* ou *E. coli*. Le vecteur peut être un polynucléotide linéaire ou,  
25 plus habituellement, un polynucléotide circulaire. Ce peut être un hybride du polynucléotide selon l'invention et d'un autre vecteur. L'autre vecteur peut être un vecteur de *E. coli*, comme pBR322, ou un vecteur de la famille pUC, R1, ColD ou RSF1010, ou un vecteur dérivé de ceux-ci

Le polynucléotide ou vecteur de l'invention peut être un plasmide. Un  
30 tel plasmide peut avoir une carte de restriction identique ou similaire aux cartes de restriction présentées sur les figures 1, 2a et 2b annexées.

Le polynucléotide ou vecteur peut avoir une taille de 1 à 20 kb, par exemple de 2 à 10 kb, de manière optimale de 3 à 7 kb.

Le polynucléotide ou vecteur peut comprendre des sites de clonage  
35 fonctionnels multiples. De tels sites de clonage comprennent généralement les séquences de reconnaissance d'enzymes de restriction. Le polynucléotide ou

vecteur peut comprendre la séquence présentée dans SEQ ID n° 1 et/ou contenir des sites de reconnaissance pour les enzymes de restriction *EcoRI*, *SacI*, *AlwNI*, *BsmI*, *BsaBI*, *BcII*, *ApaI*, *HindIII*, *SalI*, *HpaI*, *PstI*, *SphI*, *BamHI*, *Acc65I*, *EcoRV* et *BglII*. Ainsi, le polynucléotide ou vecteur peut comprendre un, plusieurs ou tous ces sites de restriction, généralement dans l'ordre indiqué sur les figures annexées.

De préférence, quand il est présent dans une bactérie telle qu'une bactérie *Propionibacterium* ou *E. coli*, le polynucléotide ou vecteur selon l'invention ne s'intègre pas dans le chromosome de la bactérie. Généralement, le polynucléotide ou vecteur ne s'intègre pas au cours de 5 générations, de préférence au cours de 20 ou 30 générations.

Le polynucléotide ou vecteur peut être un plasmide à réplication autonome qui peut rester extrachromosomique à l'intérieur d'une cellule hôte, lequel plasmide est dérivé d'un plasmide endogène de *Propionibacterium* et qui, quand il comprend un gène hétérologue (à l'hôte), est capable d'exprimer ce gène à l'intérieur de la cellule hôte. Le terme "dérivé de" signifie que le plasmide à réplication autonome comprend une séquence identique au polynucléotide de l'invention.

Le vecteur selon l'invention peut comprendre un marqueur sélectionnable qui peut être un marqueur qui confère une résistance à un antibiotique tel qu'un gène de résistance à l'ampicilline, à la kanamycine ou à la tétracycline. Le marqueur sélectionnable peut aussi être un gène de résistance à l'érythromycine qui peut provenir d'un actinomycète, comme *Saccharopolyspora erythraea*, par exemple *Saccharopolyspora erythraea* NRRL2338. D'autres marqueurs sélectionnables qui peuvent être présents dans le vecteur comprennent les gènes de résistance au chloramphenicol, à la thiostreptone, à la viomycine, à la néomycine, à l'apramycine, à l'hygromycine, à la bléomycine ou à la streptomycine.

Le vecteur peut être un vecteur d'expression qui peut comprendre un gène hétérologue (qui n'existe pas naturellement dans la cellule hôte, par exemple une cellule de *Propionibacterium*), ou un gène endogène ou homologue de la cellule hôte, par exemple de la cellule de *Propionibacterium*. Dans le vecteur d'expression, le gène à exprimer est habituellement lié de manière active à une séquence de contrôle qui est capable d'assurer l'expression du gène dans une cellule hôte.

Le terme "lié de manière active" désigne une juxtaposition dans laquelle les composants décrits sont entre eux dans une relation qui leur permet de

fonctionner de la manière prévue. Une séquence de contrôle "liée de manière active" à une séquence codante est introduite par ligature de manière que l'expression de la séquence codante soit obtenue dans des conditions compatibles avec la séquence de contrôle.

- 5                   Le gène hétérologue ou endogène peut être inséré entre les nucléotides 1 et 200 ou entre les nucléotides 1500 et 3555 de SEQ ID n° 1 ou en une position équivalente dans un polynucléotide homologue.

- De tels gènes peuvent comprendre des gènes homologues ou endogènes tels que les facteurs d'élongation, les promoteurs, les séquences ou  
10 élément de régulation, et les protéines de réplication. D'autres gènes (qui peuvent être hétérologues pour l'hôte) comprennent ceux qui codent ou qui favorisent la production de facteurs nutritionnels, d'immunomodulateurs, d'hormones, de protéines et d'enzymes (par exemple protéases, amylases, peptidases, lipases),  
15 de groupes de gènes, d'agents antimicrobiens (par exemple nisine), de substances destinées à être utilisées dans les aliments (par exemple dans les saucisses, le fromage), d'enzymes métaboliques, de vitamines (par exemple B12), d'uroporphyrinogène (III) méthyltransférase (UP III MT), de *cobA*, d'antigènes et d'agents thérapeutiques (par exemple pour vaccins). Comme on le verra, les hôtes  
20 peuvent produire une grande variété de substances, pas seulement des polypeptides, qui peuvent être le produit voulu ou qui peuvent être utilisées pour produire le produit voulu.

- Le gène hétérologue peut avoir un effet thérapeutique sur un être humain ou un animal. Un tel gène peut comprendre un antigène, provenant par  
25 exemple d'un organisme pathogène. L'hôte, tel qu'une bactérie *Propionibacterium*, qui comprend un polynucléotide ayant un tel gène hétérologue peut être utilisé comme vaccin ou dans un vaccin, et peut conférer une protection contre les agents pathogènes.

- L'antigène hétérologue peut être une protéine complète ou une partie  
30 d'une protéine contenant un épitope. L'antigène peut être issu d'une bactérie, d'un virus, d'une levure ou d'un champignon.

- La cellule hôte forme le troisième aspect de l'invention et comprend un polynucléotide ou vecteur du premier aspect ou du second aspect. La cellule hôte peut être une bactérie, par exemple de la classe des actinomycètes. La bactérie  
35 peut être une bactérie *Propionibacterium* ou *E. coli*. La bactérie

*Propionibacterium* peut être une bactérie *P. freudenreichii*, *P. jensenii*, *P. thoenii* ou *P. acidipropionici*.

Dans son quatrième aspect, l'invention concerne un procédé de production d'une cellule hôte du troisième aspect, le procédé comprenant la transformation ou la transfection d'une cellule hôte avec un polynucléotide ou vecteur du premier ou du second aspect, par exemple avec des techniques de transformation connues<sup>11</sup>.

Dans son cinquième aspect, l'invention concerne un procédé de préparation d'un polypeptide codé par le polynucléotide ou vecteur de l'invention présent dans une cellule hôte selon l'invention, qui comprend l'exposition ou la culture de la cellule hôte dans des conditions permettant l'expression du polypeptide.

Dans cet aspect, l'invention fournit donc un procédé de préparation d'un polypeptide codé par un gène donné, lequel procédé comprend la culture d'une cellule hôte transformée ou transfectée avec un vecteur d'expression comprenant le gène, dans des conditions permettant l'expression dudit polypeptide, et éventuellement la récupération du polypeptide exprimé. La cellule hôte peut être de la classe des actinomycètes, ou une bactérie Gram-positve telle qu'une bactérie *Propionibacterium* ou *E. coli*.

Les promoteurs, les initiateurs de traduction, les terminateurs de traduction, les gènes de facteurs d'élongation, l'ARN ribosomique, les gènes de résistance aux antibiotiques, les promoteurs synthétiques (par exemple conçus d'après des séquences consensus) ou les autres signaux de régulation de l'expression qui sont présents dans le polynucléotide ou vecteur peuvent être ceux qui sont compatibles avec l'expression dans la cellule hôte. De tels promoteurs comprennent les promoteurs des gènes endogènes de la cellule hôte.

Les conditions de culture peuvent être des conditions aérobies ou anaérobies, selon l'hôte. Pour un procédé de fermentation, la cellule hôte est placée dans des conditions anaérobies puis éventuellement aérobies. Les composés produits, tels qu'un polypeptide exprimé, peuvent ensuite être récupérés, par exemple à partir de la cellule hôte ou du milieu de fermentation. Le polypeptide exprimé peut être sécrété par la cellule hôte. Ou bien encore, le polypeptide peut ne pas être sécrété par la cellule hôte, auquel cas il peut être exprimé sur la surface de la cellule hôte. Ceci peut être souhaitable, par exemple, si le polypeptide comprend un antigène pour lequel on souhaite une réponse immunitaire chez l'homme ou l'animal.

Un gène homologue qui peut être présent dans le vecteur de l'invention peut être le gène *cobA*. Ainsi, une cellule hôte comprenant ce vecteur peut être capable de produire un composé tel que la vitamine B12 à partir d'un substrat, ou bien le composé peut être le produit d'une enzyme. L'invention fournit un procédé

5 de préparation de la vitamine B12 qui comprend la culture ou la fermentation d'une telle cellule hôte dans des conditions dans lesquelles le gène de la UP(III) MT est exprimé. L'enzyme produite peut être mise en contact avec un substrat approprié dans des conditions dans lesquelles le substrat est converti en vitamine B12. Ceci peut conduire à la production accrue de vitamine B12.

10 Ainsi qu'on l'a déjà indiqué, le polynucléotide selon l'invention peut comprendre un gène hétérologue qui est un gène thérapeutique. L'invention concerne donc une cellule hôte comprenant un vecteur selon l'invention qui comprend un gène thérapeutique destiné à être utilisé dans une méthode de traitement du corps humain ou animal par thérapie. Une telle cellule hôte, qui peut

15 être vivante ou morte, peut être une cellule de *Propionibacterium*.

La cellule hôte peut être formulée pour l'administration clinique en étant mélangée avec un support ou diluant pharmaceutiquement acceptable. Par exemple, elle peut être formulée pour l'administration topique, parentérale, intraveineuse, intramusculaire, sous-cutanée, orale ou transdermique. La cellule

20 hôte peut être mélangée avec tout véhicule qui est pharmaceutiquement acceptable et qui est approprié pour la voie d'administration voulue. Le support ou diluant pharmaceutiquement acceptable pour l'injection peut être par exemple une solution stérile ou isotonique comme l'eau pour injection ou le sérum physiologique.

25 La dose des cellules hôtes peut être ajustée selon différents paramètres, en particulier selon le type des cellules hôtes utilisées, l'âge, le poids et l'état du patient à traiter, le mode d'administration utilisé, la maladie à traiter et le schéma clinique nécessaire. A titre de guide, le nombre de cellules hôtes administrées, par exemple par administration orale, est de  $10^7$  à  $10^{11}$  cellules

30 hôtes par dose pour un humain adulte d'un poids corporel de 70 kg.

Les voies d'administration et les doses décrites ne constituent qu'un guide car le praticien expérimenté est capable de déterminer aisément la voie d'administration et la dose optimales pour un patient particulier quelconque et une maladie quelconque.

35 Dans un sixième aspect, l'invention fournit un polypeptide comprenant l'une des séquences d'acides aminés présentées dans SEQ ID n° 2 et 3 ou une

séquence sensiblement homologue, ou un fragment de l'une ou l'autre de ces séquences. Le polypeptide peut être un polypeptide codé par un polynucléotide du premier aspect de l'invention. En général, on préfère les séquences d'acides aminés naturelles présentées dans SEQ ID n° 2 et 3. Toutefois, les polypeptides selon  
5 l'invention comprennent les homologues des séquences naturelles, et les fragments des séquences naturelles et de leurs homologues, qui possèdent l'activité des polypeptides naturels. Une telle activité peut être d'assurer la réplication du polynucléotide selon l'invention. En particulier, un polypeptide selon l'invention peut comprendre :

- 10 (a) la protéine de SEQ ID n° 2 ou 3, ou
- (b) un homologue de celle-ci, provenant d'actinomycètes, comme *Propionibacterium freudenreichii* ou d'autres souches de *Propionibacterium*, ou
- (c) une protéine ayant une homologie d'au moins 70 % avec (a) ou (b).

15 Un homologue peut exister naturellement dans une bactérie *Propionibacterium* et peut fonctionner de manière sensiblement similaire à un polypeptide de SEQ ID n° 2 ou 3. Un tel homologue peut exister dans des actinomycètes ou des bactéries Gram-positives.

Une protéine ayant une homologie d'au moins 70 % avec les protéines  
20 de SEQ ID n° 2 ou 3 ou un homologue de celles-ci aura de préférence une homologie d'au moins 80 ou 90 % et de préférence encore d'au moins 95 %, 97 % ou 99 % avec celles-ci sur une région d'au moins 20, de préférence d'au moins 30, par exemple d'au moins 40, 60 ou 100 acides aminés contigus ou plus. Des procédés pour mesurer l'homologie des protéines sont bien connus et l'homme du  
25 métier comprendra que, dans le présent contexte, l'homologie est calculée sur la base de l'identité des acides aminés (parfois appelée "homologie stricte").

Les séquences des protéines de SEQ ID n° 2 et 3 et d'homologues peuvent donc être modifiées pour produire d'autres polypeptides dans le cadre de la présente invention.

30 Il est possible de faire des substitutions d'acides aminés, par exemple de faire 1, 2 ou 3 à 10, 20 ou 30 substitutions. En général, le polypeptide modifié conserve son activité naturelle. Il est possible de réaliser des substitutions conservatives, par exemple selon le tableau suivant. Les acides aminés de la même ligne dans la seconde colonne et de préférence de la même ligne dans la troisième  
35 colonne peuvent être substitués les uns aux autres :

Aliphatique	Non-polaire	G A P
		I L V
	Polaire - non chargé	C S T M
		N Q
	Polaire - chargé	D E
		K R
Aromatique		H F W Y

Les polypeptides selon l'invention comprennent aussi des fragments des polypeptides de longueur intégrale mentionnées ci-dessus et de leurs variantes, y compris des fragments des séquences présentées dans SEQ ID n° 2 ou 3. De tels fragments peuvent conserver l'activité naturelle du polypeptide de longueur intégrale.

Les fragments appropriés auront une taille d'au moins environ 5 acides aminés, par exemple de 10, 12, 15 ou 20 acides aminés. Les fragments de polypeptide de SEQ ID n° 2 et 3 et leurs homologues peuvent contenir une ou plusieurs (par exemple 2, 3, 5 ou 10) substitutions, délétions ou insertions, y compris des substitutions conservées.

Un polypeptide selon l'invention peut être sous une forme sensiblement isolée, et aussi sous une forme sensiblement purifiée, qui comprendra généralement le polypeptide dans une préparation dans laquelle plus de 90 %, par exemple 95 %, 98 % ou 99 % du polypeptide dans la préparation est un polypeptide de l'invention.

Un polypeptide selon l'invention peut être marqué avec un marqueur d'identification qui peut être tout marqueur approprié qui permet la détection du polypeptide. Les marqueurs appropriés comprennent les radioisotopes, par exemple  $^{125}\text{I}$ ,  $^{35}\text{S}$ , les enzymes, les anticorps, les polynucléotides et les lieux comme la biotine.

Il apparaîtra d'après la description que les cellules hôtes du troisième aspect peuvent être utilisées pour produire non seulement des protéines recombinées mais aussi d'autres composés d'intérêt, y compris des substances non protéiques telles que des substances chimiques inorganiques, en particulier des vitamines. Ainsi, dans un septième aspect, la présente invention concerne un procédé de production d'un composé, qui comprend la culture ou la fermentation de cellules hôtes du troisième aspect dans des conditions permettant la production du composé voulu. Bien que ce composé puisse être un polypeptide, par exemple



un polypeptide du sixième aspect, ce peut être également l'un des composés mentionnés ci-dessus au sujet des gènes à exprimer. Il est clair que les composés inorganiques ne sont pas exprimés par un gène mais peuvent être produits par une enzyme, ou bien le polypeptide ou l'enzyme peut aider la cellule hôte à produire les composés voulus. Ces composés peuvent être produits dans la cellule et isolés ultérieurement, par exemple après la lyse de la cellule hôte, ou bien ils peuvent franchir la paroi de la cellule hôte pour parvenir dans le milieu environnant qui peut être un milieu de fermentation, par exemple une solution aqueuse. De cette manière, les cellules hôtes peuvent être cultivées dans un milieu aqueux qui comprend des cellules et des produits nutritifs pour les cellules, par exemple des sources assimilables de carbone et/ou d'azote.

Les polypeptides ainsi produits peuvent avoir des utilisations thérapeutiques. Ils peuvent constituer des médicaments ou d'autres composés pharmacologiquement actifs, ou bien ils peuvent être antigéniques ou immunogènes, auquel cas ils peuvent être utilisés dans des vaccins.

La présente invention concerne en outre le composé produit par ce procédé, qu'il s'agisse ou pas d'un polypeptide recombiné. Les vitamines, comme la vitamine B12 (cobalamine), sont des composés spécifiquement envisagés.

Dans certains cas, il n'est pas nécessaire d'isoler le composé à partir du milieu de fermentation ou des cellules hôtes. Les cellules hôtes peuvent être utilisées elles-mêmes dans des applications particulières, par exemple dans la fabrication d'aliments tels que les saucisses, ou dans la fabrication du fromage, ou bien les cellules hôtes peuvent être incluses dans des aliments pour animaux, par exemple quand les cellules hôtes contiennent un composé destiné à être ingéré par les animaux en question. De ce fait, l'invention concerne aussi l'utilisation de ces composés ou de ces cellules hôtes dans la production d'aliments tels que les fromages et les saucisses. L'invention concerne aussi les aliments ou les aliments pour animaux comprenant des cellules hôtes ou un composé produit selon l'invention.

Dans un mode de réalisation de la présente invention que l'on préfère particulièrement, il est possible d'utiliser les cellules hôtes dans un procédé de fabrication du fromage, de sorte que la présente invention concerne également un procédé de fabrication du fromage dans lequel les micro-organismes employés sont des cellules hôtes selon l'invention. Les cellules hôtes peuvent être utilisées à la place de ou en plus d'autres bactéries, telles que les bactéries lactiques. Les bactéries propioniques sont couramment utilisées dans les procédés de fabrication

du fromage, par exemple avec des cultures mésophiles (fromage de type Maasdam) et avec des cultures thermophiles (Emmental). Les organismes mésophiles et les organismes thermophiles peuvent être responsables de l'acidification du lait ou du fromage. Ainsi, il est possible d'utiliser les cellules hôtes de l'invention non seulement pour le fromage, mais aussi pour la production d'autres produits laitiers fermentés (par exemple yaourt). Les bactéries propioniques sont employées dans la fabrication du fromage du fait de leur aptitude à convertir le lactate et les glucides en acide propionique, acide acétique et dioxyde de carbone. Les cellules hôtes selon l'invention, en particulier s'il s'agit de bactéries *Propionibacterium*, peuvent être employées car elles sont moins sensibles aux nitrates et au sel, ce qui peut permettre la réduction ou l'omission de la bactofugation du lait (employée habituellement pour réduire les niveaux de *Clostridia*).

La fermentation des cellules hôtes peut comprendre une ou deux phases ou étapes. Il peut y avoir par exemple une phase de croissance et/ou de production, ou une phase anaérobie et/ou aérobie. De préférence il y aura une phase de croissance et/ou anaérobie, et judicieusement aussi (par exemple ultérieurement) une phase de production et/ou aérobie.

Les sources de carbone et d'azote peuvent être des sources complexes ou des composés individuels. Pour le carbone, il est préférable que la source soit constituée par le glucose. Pour l'azote, les sources appropriées comprennent les extraits de levure, l'ammoniaque et les ions ammonium.

Les caractéristiques que l'on préfère pour un aspect de l'invention conviennent aussi pour un autre aspect en faisant les changements nécessaires.

L'invention sera mieux comprise à la lecture des exemples non limitatifs suivants et des dessins annexés dans lesquels :

la figure 1 est une carte de restriction du vecteur p545 selon l'invention, obtenu à partir de *P. freudenreichii* LMG 16545 (CBS 101022) ; et

les figures 2a et 2b comportent chacune deux cartes de restriction de deux vecteurs selon l'invention.

### Exemple 1

#### Criblage de souches de *Propionibacterium*

On a criblé une collection de 75 souches non pathogènes de *Propionibacterium* concernant la présence de plasmides endogènes. On a obtenu la plus grande partie des souches auprès de la collection de cultures BCCM/LMG

(Gand, Belgique), bien que l'on ait obtenu certaines souches auprès de la ATCC (Rockville, Md., USA) ou de la DSM (Braunschweig, Allemagne). On a réalisé le criblage en utilisant un protocole d'isolement de plasmides à petite échelle. Tout d'abord, on a cultivé les bactéries dans des conditions anaérobies dans du milieu MRS<sup>6</sup> pendant 48 h à 30°C. Puis on a purifié les plasmides à partir des bactéries au moyen d'un protocole d'isolement d'ADN plasmidique développé à l'origine pour *E. coli*<sup>4</sup> avec certaines modifications : on a lavé les cellules d'une culture de 5 ml dans une solution de saccharose à 25 %, tris-HCl pH 8 50 mM, on les a remises en suspension dans 250 µl de TENS (saccharose 25 % + NaCl 50 mM + tris-HCl 50 mM + EDTA 5 mM pH 8), contenant 10 mg/ml de lysozyme (Boehringer Mannheim), et on a incubé à 37°C pendant 20-30 min. Puis on a lysé les cellules bactériennes dans 500 µl de NaOH 0,2 N/SDS 1 % (incubation de 2 à 5 min sur la glace). Après avoir ajouté 400 µl de NaAc 3M pH 4,8 (5 min sur la glace) puis extrait avec du phénol/chloroforme, on a précipité l'ADN en ajoutant de l'isopropanol.

On a analysé l'ADN par électrophorèse sur des gels d'agarose à 1 % et on l'a visualisé au bromure d'éthidium. Tandis que certaines souches étaient négatives, c'est-à-dire qu'elles ne révélaient pas la présence de plasmides endogènes dans cette analyse, la plus grande partie des souches se sont révélées positives et contenaient des plasmides de grande taille ( $\geq 20$  kb). On a observé des plasmides plus petits dans six souches. Parmi celles-ci, *P. jensenii* LMG16453, *P. acidipropionici* ATCC4875, *P. acidipropionici* LMG16447 et une souche non spécifiée de *Propionibacterium* (LMG16550) contenaient un plasmide dans le domaine de taille de 6 à 10 kb. Deux souches (*P. freudenreichii* LMG16545 et *P. freudenreichii* LMG16546) présentaient un profil de plasmides identique de 2 plasmides. Un plasmide était de taille importante (non déterminée) et l'autre était plus petit, plus abondant et avait une taille estimée de 3,6 kb. On a choisi les plasmides de 3,6 kb de LMG16545 et LMG16546 pour une analyse plus poussée.

## 30 Exemple 2

### Analyse d'un plasmide endogène provenant des souches LMG16545 et LMG16546

On a isolé les plasmides de 3,6 kb des deux souches ci-dessus et on les a purifiés encore par ultracentrifugation en gradient de densité de CSCI-bromure d'éthidium<sup>11</sup>. On a établi des cartes de restriction limitées des deux préparations et il est apparu qu'elles étaient identiques<sup>11</sup>. La carte de restriction du plasmide de

3,6 kb est montrée sur la figure 1. On a obtenu les enzymes de restriction et la ligase de T4 auprès de New England Biolabs ou GIBCO BRL.

On a radiomarké et utilisé dans des expériences d'hybridation par transfert de Southern le plasmide de 3,6 kb provenant de la souche LMG16545 (que l'on appellera dans la suite p545). Les conditions d'hybridation étaient 0,2 x SSC, 65°C. Ce plasmide réagissait aussi bien avec des extraits d'ADN plasmidique de LMG16545 qu'avec des extraits d'ADN plasmidique de LMG16546, ce qui confirme la relation étroite de ces souches, tandis qu'un extrait d'ADN plasmidique provenant de *P. acidipropionici* ATCC4875, qui contient un plasmide de 6,6 kb appelé pTY1 ou pRG01<sup>8</sup>, ne réagissait pas.

On a déterminé la séquence d'ADN du plasmide p545 avec des didésoxyribonucléotides marqués avec un colorant fluorescent dans un séquenceur automatique Applied Biosystems 373A, et on l'a incluse dans la liste de séquences annexée, sous forme de SEQ ID n° 1. On a réalisé l'analyse de la séquence sur l'ADN plasmidique que l'on avait linéarisé avec *EcoRI* et inséré dans l'ADN de pBluescript SKII+ digéré avec *EcoRI* (Stratagene, La Jolla, Ca., USA). L'analyse assistée par ordinateur de la séquence ainsi obtenue à l'aide de la technique BLAST<sup>1</sup> a révélé des homologues avec des protéines impliquées dans la réplication de plasmides provenant de plusieurs organismes riches en GC (par exemple *repA* et *repB* codés par pAL5000 provenant de *Mycobacterium fortuitum*<sup>8,14</sup> présentent 28 à 30 % d'identité et 34 à 38 % de similitude avec les protéines de réplication supposées respectives provenant du plasmide p545 ; pXZ10142 provenant de *Corynebacterium glutamicum* [numéro d'ordre PIR S32701] est un autre exemple de plasmides codant des protéines de réplication homologues des protéines de réplication supposées de p545). Les résultats des comparaisons au moyen de bases de données avec des séquences homologues sont présentés dans les exemples 7 et 8.

### Exemple 3

#### 30 Construction de vecteurs navettes de *E. coli*/*Propionibacterium*

On a digéré le plasmide pBR322 de *E. coli* avec *EcoRI* et *AvaI* et on a remplacé par un ADN duplex synthétique le plus petit fragment ainsi produit, qui mesure 1,4 kb et qui comprend le gène de résistance à la tétracycline. On a conçu l'ADN duplex synthétique de manière à relier les extrémités *EcoRI* et *AvaI* et à  
35 fournir un certain nombre de sites de restriction uniques :

5' -

AATTCAAGCTTGTGCGACGTTAACCTGCAGGCATGCGGATCCGGTACCG  
ATATCAGATCT - 3' (SEQ ID n° 4)

3' -

5 GTTCGAACAGCTGCAATTGGACGTCCGTACGCCTAGGCCATGGCTATA  
GTCTAGAAGCC - 5' (SEQ ID n° 5)

Les sites de restriction suivants ont été fournis de cette manière :

*EcoRI* (rétabli), *HindIII*, *Sall*, *HpaI*, *PstI*, *SphI*, *BamHI*, *Acc65I*, *EcoRV*, *BglIII*  
(*AvaI* n'est pas rétabli).

10 On a ligaturé cet ADN synthétique avec le grand fragment et on a  
transféré de nouveau le mélange de ligature dans *E. coli* (en utilisant la ligase de  
T4). On a obtenu ainsi un plasmide ayant la composition prévue (pBR322ΔI). On  
peut utiliser le site de clonage multiple pour introduire un marqueur de sélection  
ainsi que l'ADN du plasmide p545.

15 On va décrire à titre d'exemple la construction d'un plasmide navette  
de *E. coli/Propionibacterium* qui confère une résistance à l'érythromycine.

On a inséré dans pBR322ΔI linéarisé avec *Acc65I* un fragment *Acc65I*  
de 1,7 kb provenant du groupe de biosynthèse de l'érythromycine de  
*Saccharopolyspora erythraea* NRRL2338 et contenant le gène de résistance à  
20 l'érythromycine<sup>15,2</sup> puis on a linéarisé avec *EcoRV* la construction ainsi obtenue,  
appelée pBRES, et on l'a ligaturée avec l'ADN de p545 que l'on avait digéré avec  
*BsaBI*. On a constaté que les transformants *E. coli* contenaient un vecteur ayant  
l'insert correct, dans les deux orientations. On a appelé les vecteurs plasmidiques  
résultants pBRESP36B1 et pBRESP36B2 (figures 2a et 2b).

25 On a également obtenu des constructions de vecteurs plasmidiques  
avec l'ADN de p545 linéarisé dans un autre site de restriction situé à l'extérieur de  
la région de réplication supposée, c'est-à-dire *AlwNI*. Pour cette construction, il a  
été nécessaire de munir le vecteur pBRES d'un site de clonage approprié. On a  
conçu un adaptateur consistant en deux oligonucléotides complémentaires de  
30 composition suivante (SEQ ID n° 6 et 7) :

5' GTACCGGCCGCTGCGGCCAAGCTT 3'

5' GATCAAGCTTGGCCGCAGCGGCCG 3'

L'hybridation (annelage) de ces oligonucléotides crée un fragment  
d'ADN double brin avec des extrémités cohésives *Acc65I* et *BglIII*, respectivement,  
35 qui contient en outre un site de restriction *SfiI* interne, qui donne des extrémités  
compatibles avec le plasmide p545 digéré avec *AlwNI*. On a cloné cet adaptateur

dans pBRES entre le site *Bgl*II et le site *Acc*65I proximal. Puis on a digéré avec *Sfi*I le vecteur pBRES-*Sfi*I ainsi obtenu et on l'a ligaturé avec p545 digéré avec *Alw*NI. La transformation de *E. coli* a donné des transformants ayant le vecteur correct, comme l'a confirmé l'analyse avec des enzymes de restriction. On a appelé le vecteur obtenu pBRESP36A (figure 2).

#### Exemple 4

#### Transformation de *propionibacterium* avec des vecteurs navettes de *E. coli*/Propionibacterium

On va décrire la transformation de la souche de *Propionibacterium freudenreichii* ATCC6207 avec pBRESP36B1.

On cultive les cellules bactériennes dans du milieu SLB (bouillon au lactate de sodium<sup>17</sup>) à 30°C jusqu'à la phase de croissance stationnaire, puis on dilue 1:50 dans du milieu SLB frais. Après environ 20 h d'incubation à 30°C, on récolte les cellules (qui sont maintenant dans la phase de croissance exponentielle) et on les lave dans du saccharose 0,5M froid. Puis on les lave une fois dans du tampon d'électroporation consistant en saccharose 0,5M, tamponné par de l'acétate de potassium 1 mM, pH 5,5, et enfin on les remet en suspension dans ce tampon d'électroporation dans un volume sensiblement égal au centième du volume de culture initial. On maintient les cellules sur la glace pendant la totalité du protocole.

Pour l'électroporation (appareil de BIORAD), on mélange 80 à 100 µl de la suspension de cellules avec environ 1 µg d'ADN (ou des quantités plus faibles), dans une cuvette d'électroporation de 1 ou 2 mm refroidie, et on applique une impulsion électrique. On a constaté que les conditions optimales pour les impulsions étaient 25 kV/cm à une résistance de 200 Ω et à une capacité de 25 µF. Toutefois, des tensions plus basses et plus élevées (également à 100 Ω) produisent aussi des transformants.

Immédiatement après l'impulsion, on ajoute 900 µl de SLB froid contenant 0,5 M de saccharose à la suspension de cellules que l'on incube ensuite pendant 2,5 à 3 h à 30°C avant d'étaler des dilutions appropriées sur des boîtes de SLB/gélose contenant 0,5 M de saccharose et 10 µg/ml d'érythromycine. Après une période d'incubation de 5 à 7 jours à 30°C dans des conditions anaérobies, on détecte des transformants.

L'ADN isolé de *E. coli* DH5α (Promega) donne un rendement de transformation de 20 à 30 transformants par µg d'ADN. On obtient un rendement

10 à 100 fois supérieur quand on isole l'ADN de *E. coli* JM110 (souche dam<sup>-</sup>, dcm<sup>-</sup>). On a réalisé la transformation de *E. coli* selon les instructions de BIORAD.

Les transformants contenaient les vecteurs authentiques indiscernables de l'ADN plasmidique initial utilisé pour la transformation de ATCC6207. On a  
5 mis ceci en évidence par analyse par des enzymes de restriction d'ADN plasmidique isolé à partir des transformants par le protocole d'isolement d'ADN plasmidique à petite échelle évoqué ci-dessus.

Les vecteurs étaient présents exclusivement sous forme de plasmides à  
réplication autonome. L'hybridation par transfert de Southern<sup>13</sup> avec des isolats  
10 d'ADN total a montré que l'ADN chromosomique ne s'hybridait pas avec l'ADN vecteur utilisé comme sonde, ce qui indique qu'il ne se produit pas d'intégration chromosomique de l'ADN plasmidique.

La transformation a été réussie également avec les vecteurs  
pBRESP36B2 et pBRESP36A, ce qui indique que la fonctionnalité des vecteurs  
15 est indépendante de l'orientation de p545 ou du site de clonage utilisé. Dans ce cas également, on a confirmé l'authenticité des vecteurs.

En outre, la transformation de la souche de *P. freudenreichii*  
ATCC6207 avec de l'ADN isolé à partir d'un transformant de *Propionibacterium* a  
eu pour conséquence un rendement de transformation augmenté de 10<sup>5</sup> à 10<sup>6</sup> fois  
20 par rapport à celui obtenu avec l'ADN isolé à partir de *E. coli* DH5 $\alpha$ .

La transformation d'une autre souche de *P. freudenreichii*, LMG16545  
(la souche à partir de laquelle on a obtenu le plasmide p545), a donné un  
rendement de transformation comparable à celui de la souche ATCC.

Les résultats des transformations, et l'effet sur la production de la  
25 vitamine B12, sont présentés dans le tableau suivant.

Huit transformants parmi 10 ont donné une teneur en vitamine B12  
supérieure de 50 % à celle de la souche témoin.

Souche	Plasmide transformé	Vitamine B12 (mg/g de matière sèche)
COB1	PBRES36COB	0,57
COB2	pBRES36COB	0,67
COB3	pBRES36COB	0,83
COB4	pBRES36COB	0,68
COB5	pBRES36COB	0,69
COB6	pBRES36COB	0,61
COB7	pBRES36COB	0,53
COB8	pBRES36COB	0,64
COB9	pBRES36COB	0,50
COB10	pBRES36COB	0,74
recATCC6207	pBRESP36B2	0,54

### Exemple 5

#### Construction d'un vecteur plasmidique contenant le gène *cobA*

5 On va décrire la construction et l'application d'un vecteur plasmidique pour augmenter le niveau de synthèse de la vitamine B12 (cobalamine) dans la souche de *P. freudenreichii* ATCC6207.

On a formé la région promotrice du gène conférant une résistance à l'érythromycine dans *Saccharopolyspora erythraea*<sup>2,3</sup>, par PCR en utilisant les amorces suivantes (SEQ ID n° 8 et 9) :

amorce directe : (5' - 3')

AAACTGCAGCTGCTGGCTTGCGCCCGATGCTAGTC

amorce inverse : (5' - 3')

AAACTGCAGCAGCTGGGCAGGCCGCTGGACGGCCTGCCCTCGAGCTCG

15 TCTAGAATGTGCTGCCGATCCTGGTTGC

Le fragment de PCR ainsi formé contient un site *AlwNI* à l'extrémité 5' suivi par la région promotrice authentique et les 19 premiers acides aminés de la région codante du gène de résistance à l'érythromycine, pour garantir une initiation appropriée de la transcription et de la traduction. A l'extrémité 3', on dispose des sites *XbaI* et *XhoI* (pour l'insertion du gène *cobA* dans une étape ultérieure), une séquence terminatrice présente en aval du gène de résistance à l'érythromycine, et un site *AlwNI*.



On a digéré le produit de PCR avec *AlwNI* et on l'a ligaturé avec pBRESP36B2, digéré partiellement avec *AlwNI*. Parmi les deux sites *AlwNI* présents dans pBRESP36B2, seul celui qui est présent dans la partie spécifique de p545 du vecteur reçoit le fragment. On a obtenu des transformants de *E. coli* contenant la construction prévue, appelée pBRES36pEt. On a utilisé ce vecteur pour des constructions ultérieures que l'on décrira dans la suite.

On a formé la séquence codante de *cobA*, le gène codant l'uroporphyrinogène III méthyltransférase, par PCR à partir de la souche de *Propionibacterium freudenreichii* ATCC6207, en utilisant les amorces suivantes (SEQ ID n° 10 et 11) :

directe : (5' - 3')

CTAGTCTAGACACCGATGAGGAAACCCGATGA

inverse : (5' - 3')

CCCAAGCTTCTCGAGTCAGTGGTCGCTGGGCGCGCG

Le gène *cobA* ainsi amplifié porte un site *XbaI* dans la région codant l'extrémité N-terminale, et des sites *HindIII* et *XhoI* au niveau de la région codant l'extrémité C-terminale.

On a confirmé la fonctionnalité de ce gène *cobA* en clonant le produit de PCR sous forme d'un fragment *XbaI-HindIII* dans pUC18, puis en transformant la souche de *E. coli* JM109. Les transformants qui ont un gène *cobA* fonctionnel présentent une fluorescence rouge clair quand ils sont éclairés avec de la lumière UV. On a digéré avec *XbaI* et *XhoI* l'ADN plasmidique isolé à partir d'un tel transformant, on l'a ligaturé avec l'ADN de pBRESP36B2 digéré de la même manière et on l'a utilisé pour transformer *E. coli*. On a analysé par digestion par des enzymes de restriction et électrophorèse en gel l'ADN provenant de plusieurs transformants. On a constaté que les transformants portaient l'insert correct dans le vecteur d'expression. On a appelé ce nouveau vecteur pBRES36COB. Puis on a transféré ce vecteur dans *P. freudenreichii* ATCC6207 en suivant le protocole décrit ci-dessus. On a analysé dix des transformants obtenus et on a constaté qu'ils contenaient le vecteur pBRES36COB qui, de nouveau, était indiscernable du vecteur initial utilisé pour la transformation, comme le montre l'analyse des produits de digestion des enzymes de restriction. Dans ces dix transformants, on a déterminé le niveau de synthèse de la vitamine B12 de la manière suivante :

On a inoculé des cultures congelées des transformants de *Propionibacterium* 1 à 10, ainsi qu'une souche témoin contenant seulement le vecteur plasmidique pBRESP36B2, dans des flacons de 100 ml contenant 50 ml de

milieu BHI (infusion à base de cervelle et de cœur) (Difco) et on a incubé pendant 72 h à 28°C sans secouer. On a transféré 4 ml de cette préculture dans 200 ml d'un milieu de production consistant en 15 g/l d'extrait de levure Difco, 30 g/l de lactate de Na, 0,5 g/l de  $\text{KH}_2\text{PO}_4$ , 0,01 g/l de  $\text{MnSO}_4$  et 0,005 g/l de  $\text{CoCl}_2$  dans un flacon à secousses de 500 ml et on a incubé à 28°C pendant 56 h sans secouer, puis pendant 48 h dans une secoueuse rotative New Brunswick à 200 tours par minute.

On a mesuré les titres de vitamine B12 en utilisant le procédé par chromatographie liquide à haute performance publié par Blanche (Analytical Biochemistry, 1990). Parmi les dix transformants, neuf présentaient une production de vitamine B12 supérieure d'environ 25 % à la souche témoin.

#### Exemple 6

##### Stabilité des plasmides

Les trois vecteurs navettes pBRESP36A, pBRESP36B1 et pBRESP36B2 sont tous maintenus de manière stable pendant 30 générations de culture des transformants respectifs : on n'a observé aucune perte de résistance à l'érythromycine, déterminée par des dénombrements de viabilité sur des boîtes de gélose sélective (contenant de l'érythromycine) et sur des boîtes de gélose non sélective. On a établi la stabilité structurale du plasmide dans la population de transformants au bout de 30 générations en isolant l'ADN plasmidique et en le caractérisant par cartographie de restriction comme décrit ci-dessus : on a observé uniquement des fragments de restriction semblables à ceux du plasmide authentique.

#### Exemple 7

Analyse de l'homologie de séquence au moyen d'une base de données pour le polypeptide prévu codé par le premier cadre de lecture ouvert (SEQ ID n° 2)

MDSFETLFPESWLPRKPLASAEKSGAYRHVTRQRALELPYIEANPLVMQSL  
VITDRDASDADWAADLAGLPSYVSMNRVTTTGHIVYALKNPVCLTDA  
ARRRPINLLARVEQGLCDVLGGDASYGHRITKNPLSTAHATLWGPADALY  
ELRALAHTLDEIHALPEAGNPRRNVTSTVGRNVTLFDTTMWAYRAVR  
HSWGGPVAEWEHTVFEHIHLLNETIAD

On a aligné la séquence de 227 acides aminés ci-dessus (ORF1) et on l'a comparée avec plusieurs autres séquences protéiques (cible NBRF-PIR, publication PIR R52.0, mars 1997, coupure 45, KTUP:2).

5 Avec une protéine issue du plasmide de *Mycobacterium fortuitum* pAL 5 000 (JS0052), on a constaté un appariement de 37,1 % sur 194 acides aminés (INIT 167 292). Avec une protéine provenant de *Corynebacterium glutamicum* (S32701), on a trouvé un appariement de 32,0 % sur 125 acides aminés (INIT 125, 116). On a trouvé un appariement de 29,9 % sur 221 acides aminés (INIT 86, 259) avec la protéine ColE2 de *E. coli* (SO4455). On a trouvé  
10 exactement le même appariement sur le même nombre d'acides aminés pour la protéine ColE3 provenant elle aussi de *E. coli* (SO4456).

#### Exemple 8

Analyse de l'homologie de séquence par base de données pour le polypeptide  
15 prévu codé par le second cadre de lecture (SEQ ID n° 3)

MTTRERLPRN GYSIAAAAKK LGVSESTVKR WTSEPREEFV ARVAARHARI  
RELRSEGQSM RAIAAEVGVS VGTVHYALNK NRTDA

20 On a aligné aussi la seconde protéine (OFR2) et on l'a comparée avec une autre protéine en utilisant les mêmes paramètres et le même logiciel que ceux décrits dans l'exemple 7. Toutefois, cette séquence a une longueur de 85 acides aminés seulement.

On a comparé la séquence ORF2 avec une protéine provenant de  
25 *Mycobacterium fortuitum* (S32702) et on a trouvé un appariement de 53,3 % sur 75 acides aminés (INIT 207, 207).

#### Exemple 9

Analyse fonctionnelle du plasmide p545

30 Pour améliorer encore le système de vecteur, on a délimité plus précisément les fonctions plasmidiques essentielles pour la réplication et la stabilité en délétant de grandes régions du plasmide p545 initial. Le résultat obtenu, un vecteur de clonage plus petit, permet d'utiliser le système de vecteur basé sur p545 pour cloner de plus grands fragments d'ADN.

35 A cette fin, on a digéré le vecteur pBRESP36A (figure 2) avec *Sst*II et *Bcl*II, ce qui a donné un fragment de 1,7 kb et un fragment de 6,5 kb. On a

remplacé le fragment de 1,7 kb, en fait le fragment *AlwNI-BsII* de 1,6 kb du plasmide p545, par un ADN duplex synthétique composé de SEQ ID n° 12 et SEQ ID n° 13 comportant des extrémités compatibles avec *SstII* et *BclI* et un certain nombre de sites de restriction uniques.

5

SEQ ID n° 12 5' GGAGATCTAGATCGATATCTCGAG 3'

SEQ ID n° 13 5' GATCCTCGAGATATCGATCTAGATCTCCGC 3'

On a obtenu de cette manière les sites de restriction suivants :

10

*SstII* (rétabli), *BglII*, *XbaI*, *Clal*, *EcoRV*, *XhoI*, (*BclI* n'est pas rétabli).

On a transféré le mélange de ligature dans *E. coli* et on a sélectionné des transformants contenant un vecteur ayant la composition prévue. On appelé ce vecteur pBRESAΔS-B. La transformation réussie subséquente de la souche de *P. freudenreichii* ATCC6207 avec ce vecteur indiquait que la région de 1,6 kb entre *AlwNI* et *BclI* dans p545 n'est essentielle pour la réplication du plasmide.

15

On a fait une délétion supplémentaire en retirant les 240 pb correspondant à la région située entre *SalI* et *BclI* dans le plasmide p545 en digérant pBRESAΔS-B avec *SalI-SstI* et avec *SstI-XhoI*, respectivement, et en isolant le fragment *SalI-SstI* de 1,3 kb et le fragment *SstI-XhoI* de 6,6 kb. On a ligaturé les fragments et on a transféré le mélange de ligature dans *P. freudenreichii* ATCC6207, ce qui a donné de nombreux transformants. On a isolé la construction nouvellement obtenue, appelée pBRESAΔS-S, et on a confirmé sa structure par cartographie de restriction.

20

Ainsi, toutes les informations essentielles pour la réplication du plasmide p545 sont situées sur un fragment de 1,7 kb délimité par les sites de restriction *SalI* et *AlwNI* et englobant les protéines de réplication prévues codées par ORF1 et ORF2, et il est possible de déléter ces 1,8 kb sans perturber de manière manifeste la réplication ou la stabilité du plasmide.

25

### 30 Exemple 10

#### Expression du gène de résistance au chloramphénicol de *Corynebacterium*

On a établi qu'un gène de résistance au chloramphénicol (*cml*) provenant de *Corynebacterium*<sup>18</sup> code une protéine d'export du chloramphénicol. On a inséré ce gène dans le vecteur navette de *Propionibacterium-E. coli* pBRESP36B2. On a digéré ce vecteur avec *BglII* et *HindIII*, et avec *BglII* et *HpaI*

35

respectivement. On a isolé les fragments *Bgl*III-*Hind*III de 2,9 kb et *Bgl*III-*Hpa*I de 5,2 kb.

On a obtenu le fragment contenant le gène *cml*, y compris son propre promoteur, par digestion avec *Pvu*II et *Hind*III, et on a isolé le fragment de 3,3 kb  
5 contenant le gène. On a ligaturé les deux fragments spécifiques du vecteur et le fragment *cml* : les extrémités *Pvu*II et *Hpa*I sont cohésives, ce qui a permis l'insertion du gène *cml* et le rétablissement du gène *ermE* du vecteur parental. On a transféré le mélange de ligature dans *E. coli* et on a choisi un transformant dans lequel le vecteur contenait l'insert *cml* correct. On a appelé ce vecteur  
10 pBRESBCM.

La transformation de *P. freudenreichii* ATCC6207 avec ce vecteur, et la sélection sur des boîtes contenant 10 µg/ml d'érythromycine ou 5 µg/ml de chloramphénicol ont donné des colonies résistant à l'érythromycine et au chloramphénicol, respectivement, ce qui indique que, mis à part le gène de  
15 résistance à l'érythromycine (présenté précédemment avec les vecteurs navettes de *Propionibacterium-E. coli*), le gène de résistance au chloramphénicol est exprimé également dans *Propionibacterium*. On a pu cultiver les transformants dans un milieu liquide contenant jusqu'à 20 µg/ml de chloramphénicol.

## 20 Exemple 11

### Expression du gène de la lipase (*gehA*) provenant de *P. acnes*

Pour illustrer le clonage et l'expression efficaces d'une protéine extra-cellulaire à l'aide du présent système de vecteur, on a utilisé un gène de lipase *gehA* de *P. acnes*<sup>19</sup>. On a digéré avec *Xho*I le vecteur pUL6001, contenant *gehA*  
25 sur un fragment *Xho*I, et on a isolé le fragment de 2,75 kb contenant le gène. On a linéarisé le vecteur pBRESAΔS-B (provenant de l'exemple 9) avec *Xho*I, et on a déphosphorylé les extrémités avec la phosphatase intestinale de veau pour éviter l'autoligature. On a ligaturé le vecteur linéarisé et le fragment contenant *gehA* et on a transféré le mélange de ligature dans *E. coli*. On a analysé les transformants  
30 par analyse de restriction en ce qui concerne la présence du plasmide recombiné correct, appelé pBRESALIP. Puis, on a transféré ce plasmide dans la souche de *P. freudenreichii* ATCC6207. On a criblé les transformants concernant l'expression du gène de la lipase en utilisant des boîtes de gélose contenant de la tributyrine comme indicateur de l'expression accrue de la lipase. Les transformants de *P.*  
35 *freudenreichii* contenant pBRESALIP, présentaient des tailles de halo

sensiblement accrues dans cet essai par rapport aux souches non transformées ou aux souches transformée avec le vecteur parental.

### Références

- 5 Altschul et al (1990) J. Mol. Biol. 215, 103  
Bibb et al., (1985) Gene 38, 215  
Bibb et al., (1994) Mol. Microbiol. 14(3), 533  
Birnboim et Doly (1979) Nucleic Acids Res. 7, 1513  
Blanche (1990) Anal. Biochem. 189, 24  
10 DeMan et al (1960) J. Appl. Bacteriol. 36, 130  
Labidi et al (1992) Plasmid 27, 130  
Rehberger et Glatz (1990) Appl. Environ. Microbiol. 59, 83  
Rossi et al (1996) Res. Microbiol. 147, 133  
Sambrook et al (1989) Molecular Cloning, Cold Spring  
15 Harbor Laboratory Press  
Sattler et al (1995) J. Bact. 177, 1564  
Southern (1975) J. Mol. Biol. 98, 503  
Stolt et Stocker (1996) Microbiol. 142, 2795  
Thompson et al (1982) Gene 20, 51  
20 Uchijama et Weisblum (1985) Gene 38, 103  
de Vries et al (1972) J. Gen. Microbiol. 71, 515

### RESUME DES SEQUENCES

1. Séquence d'ADN du plasmide LMG 16545 (CBS 101022), 3,6 kb.  
25 2. Acides aminés de la protéine de ORF1 (303 résidus, bases 273-1184).  
3. Acides aminés de la protéine de ORF2 (85 résidus, bases 1181-1438).  
4-13. Amorces d'ADN/oligonucléotides.

### LISTE DE SEQUENCES

#### 30 (1) INFORMATIONS GENERALES :

##### (i) DEMANDEUR :

(A) NOM : Gist-Brocades B.V.

(B) RUE : Wateringseweg 1

(C) VILLE : Delft

35 (E) PAYS : Pays-Bas

(F) CODE POSTAL (ZIP) : 2600 MA

(ii) TITRE DE L'INVENTION : Polynucléotide, vecteur et cellule hôte  
le contenant, polypeptide  
correspondant, procédé pour leur  
production et leur utilisation.

5 (iii) NOMBRE DE SEQUENCES : 13

(iv) FORME LISIBLE PAR ORDINATEUR :

(A) TYPE DE SUPPORT : Disquette

(B) ORDINATEUR : IBM PC Compatible

(C) SYSTEME D'EXPLOITATION : PC-DOS/MS-DOS

10 (D) LOGICIEL : PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATIONS POUR SEQ ID N° 1 :

(i) CARACTERISTIQUES DE LA SEQUENCE :

(A) LONGUEUR : 3555 paires de bases

15 (B) TYPE : Acide nucléique

(C) NOMBRE DE BRINS : Deux

(D) TOPOLOGIE : Circulaire

(ii) TYPE DE MOLECULE : ADN (génomique)

(iii) HYPOTHETIQUE : NON

20 (iii) ANTI-SENSE : NON

(vi) SOURCE INITIALE :

(A) ORGANISME : Propionibacterium freudenreichii

(C) ISOLAT INDIVIDUEL : CBS101022 LMG16545

(ix) CARACTERISTIQUES :

25 (A) NOM/CLE : CDS

(B) POSITION : 273..1184

(D) AUTRES INFORMATIONS : /gène= "ORF1"

(ix) CARACTERISTIQUES

(A) NOM/CLE : CDS

30 (B) POSITION : 1181..1438

(D) AUTRES INFORMATIONS : /gène= "ORF2"

(xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 1 :

35 GTCGACCCTG ACAGCCGGCG AGCAGTTCAG GCGAAGATCG CACAGCTGCG CGAGGAACTA  
60

GCCGCAATGC CCGAACACGC CCCAGCCATC CCTTGGAGCA GGTGGCAGCG TCAGGGGAGT  
120

CGGGGGATGT TTGGCAGGGG ATGTGGAAAG AGAGTTCGCT TTGCTCACAT GGCTCAACCG  
180

5 GGTAAC TAAC TGATATGGGG TCTTCGTCGC CCACTTTGAA CACGCCGAGG AATGGACCAC  
240

GCTGAACGTG ACTCGCATGC TTCACTGCAT GT ATG GAT TCG TTC GAG ACG TTG  
293

10 Met Asp Ser Phe Glu Thr Leu  
1 5

TTC CCT GAG AGC TGG CTG CCA CGC AAG CCG CTG GCG TCA GCC GAG AAG  
341

15 Phe Pro Glu Ser Trp Leu Pro Arg Lys Pro Leu Ala Ser Ala Glu Lys  
10 15 20

TCT GGG GCG TAC CGG CAC GTG ACT CGG CAG AGG GCG CTG GAG CTG CCT  
389

20 Ser Gly Ala Tyr Arg His Val Thr Arg Gln Arg Ala Leu Glu Leu Pro  
25 30 35

TAC ATC GAA GCG AAC CCG TTG GTC ATG CAG TCC TTG GTC ATC ACC GAT  
437

25 Tyr Ile Glu Ala Asn Pro Leu Val Met Gln Ser Leu Val Ile Thr Asp  
40 45 50 55

CGA GAT GCT TCG GAT GCT GAC TGG GCC GCA GAC CTC GCT GGG CTG CCT  
485

30 Arg Asp Ala Ser Asp Ala Asp Trp Ala Ala Asp Leu Ala Gly Leu Pro  
60 65 70

TCA CCG TCC TAC GTG TCC ATG AAC CGT GTC ACG ACC ACC GGA CAC ATC  
533

35 Ser Pro Ser Tyr Val Ser Met Asn Arg Val Thr Thr Thr Gly His Ile  
75 80 85

GTC TAT GCC TTG AAG AAC CCT GTG TGT CTG ACC GAT GCC GCG CGG CGA  
581

40 Val Tyr Ala Leu Lys Asn Pro Val Cys Leu Thr Asp Ala Ala Arg Arg  
90 95 100

CGG CCT ATC AAC CTG CTC GCC CGC GTC GAG CAG GGC CTA TGC GAC GTT  
629

45 Arg Pro Ile Asn Leu Leu Ala Arg Val Glu Gln Gly Leu Cys Asp Val  
105 110 115

CTC GGC GGC GAT GCA TCC TAC GGG CAC CGG ATC ACA AAG AAC CCG CTC  
677

50 Leu Gly Gly Asp Ala Ser Tyr Gly His Arg Ile Thr Lys Asn Pro Leu  
120 125 130 135



AGC ACC GCC CAT GCG ACC CTC TGG GGC CCC GCA GAC GCG CTC TAC GAG  
 725  
 Ser Thr Ala His Ala Thr Leu Trp Gly Pro Ala Asp Ala Leu Tyr Glu  
 140 145 150  
 5  
 CTG CGC GCC CTC GCA CAC ACC CTC GAC GAG ATC CAC GCA CTG CCG GAG  
 773  
 Leu Arg Ala Leu Ala His Thr Leu Asp Glu Ile His Ala Leu Pro Glu  
 155 160 165  
 10  
 GCA GGG AAC CCG CGT CGC AAC GTC ACC CGA TCA ACG GTC GGC CGC AAC  
 821  
 Ala Gly Asn Pro Arg Arg Asn Val Thr Arg Ser Thr Val Gly Arg Asn  
 170 175 180  
 15  
 GTC ACC CTG TTC GAC ACC ACC CGC ATG TGG GCA TAC CGG GCC GTC CGG  
 869  
 Val Thr Leu Phe Asp Thr Thr Arg Met Trp Ala Tyr Arg Ala Val Arg  
 185 190 195  
 20  
 CAC TCC TGG GGC GGC CCG GTC GCC GAA TGG GAG CAC ACC GTA TTC GAG  
 917  
 His Ser Trp Gly Gly Pro Val Ala Glu Trp Glu His Thr Val Phe Glu  
 200 205 210 215  
 25  
 CAC ATC CAC CTA CTG AAC GAG ACG ATC ATC GCC GAC GAA TTC GCC ACA  
 965  
 His Ile His Leu Leu Asn Glu Thr Ile Ile Ala Asp Glu Phe Ala Thr  
 220 225 230  
 30  
 GGC CCC CTC GGC TTG AAC GAA CTT AAG CAC TTA TCT CGA TCC ATT TCC  
 1013  
 Gly Pro Leu Gly Leu Asn Glu Leu Lys His Leu Ser Arg Ser Ile Ser  
 235 240 245  
 35  
 CGA TGG GTC TGG CGC AAC TTC ACC CCC GAA ACC TTC CGC GCA CGC CAG  
 1061  
 Arg Trp Val Trp Arg Asn Phe Thr Pro Glu Thr Phe Arg Ala Arg Gln  
 250 255 260  
 40  
 AAA GCG ATC AGC CTC CGT GGA GCA TCC AAA GGC GGC AAA GAA GGC GGC  
 1109  
 Lys Ala Ile Ser Leu Arg Gly Ala Ser Lys Gly Gly Lys Glu Gly Gly  
 265 270 275  
 45  
 CAC AAA GGC GGC ATT GCC AGT GGC GCA TCA CGG CGC GCC CAT ACC CGT  
 1157  
 His Lys Gly Gly Ile Ala Ser Gly Ala Ser Arg Arg Ala His Thr Arg  
 280 285 290 295  
 50  
 CAA CAG TTC TTG GAG GGT CTC TCA TGACCACACG TGAACGTCTC CCCC GCAACG  
 1211  
 Gln Gln Phe Leu Glu Gly Leu Ser  
 300

GCTACAGCAT CGCCGCTGCT GCGAAAAAGC TCGGTGTCTC CGAGTCCACC GTCAAGCGGT  
1271

5 GGACTTCCGA GCCACGCGAG GAGTTCGTGG CCCGCGTTGC CGCACGCCAC GCGCGGATTC  
1331

GTGAGCTCCG CTCGGAGGGT CAGAGCATGC GTGCGATTGC TGCCGAGGTC GGGGTTTCCG  
1391

10 TGGGCACCGT GCACTACGCG CTGAACAAGA ATCGAACTGA CGCATGACCG TAACGCCGCA  
1451

CGATGAGCAT TTTCTTGATC GTGCACCGCT TGGCACTACG TTCGCGTGCG GTTGACAGT  
1511

15 GCGCGCCACG TTCTTATCCT GCGGCCATTG TGGCTACAGC CAATGGGGGG CATCAGCAAC  
1571

GGACGTTGAA CCCGGTGGGC AAGTGTTACT CAGGGGGACA TGCCCAGTCT GCGGCGCTCG  
20 1631

GATTGACGGT ATGGCAGTCG TGCATGCGGC CCCACCGTCA AACTCATTCA GGTATCAGTG  
1691

25 AGAACCCTCA TGGCACCCCC TCGTGACACG TTCTCGTTGC GATCAGCTGC TGTGCGTGCG  
1751

GGCGTGAGCG TTTCTACGCT GCGGCGCAGG AAATCAGAGC TTGAGGCTGC CGGAGCGACG  
1811

30 GTAGACCCGT CCGGTTGGGT GGTGCCACTG CGTGCACTCA AGGTCGTTTT TGGGGTGTC  
1871

GATGAGACCT CGAATGCGCC CGGTCATGAC GCTGAGTTAG TGGCGCAGCT GCGCTCTGAG  
35 1931

AACGAGTTTT TACGGCGTCA GGTCGAGCAG CAGGCGCGCA CGATCGAACG GCAGGCTGAG  
1991

40 GCACACGCGG TGGTCTCAGC GCAGCTCACA CGGGTTGGCC AGCTTGAGGC CGGCGACGCA  
2051

GCAGCACCGA CACTGGCACC CGTTGAAAGG CCGGCTCCGC GACGGCGGTG GTGGCAGCGT  
2111

45 CGGTAGCGGT CAGGATCGCT CTGGCGTGAC GAGTGTGTCT GGCAGTGCGA ACAGTTGCTC  
2171

GACCAGTGGC AGCAGAAGCG AGATCGCTGC GTGGTGCTGT TCCTCGGTCA GTTCGTCGAG  
50 2231

GACTGGCGGG TCTTGCTGCG TCCAGCCGAT CGCCTCGGCG GCCAAGGTCA GTTCCAAGCT  
2291

GTGCCAACGC ACACGCCCCCT CGGCTGACAG CTGAGTCTCG AACTGTGCAA CTGGACCGGC  
2351

5 CGGAAGATGC ACGTTGCCGA GGTCTGAGT GCCAAGCGC ACGTCAAAGA GTGCTGCTTC  
2411

GTAGCCGCGC AGAAATGGCA GTGCTCGGTC GATTCTGGATC GGCCTGCCCCA GGTACATTCC  
2471

10 GGGCCGCTTG ATGAACGCCT CCGCGTAGAA GCGCACCGTT CTCGGCCCCG CCTCGTGATC  
2531

TGTCACTGTG CACGCTCCTC TCGATGGTTC TCGACGCTAC CGGAGACCAC CGACGTTTAT  
2591

15 GCCCAGCGCA GCGACCTGAA AGGACCAAGC CGAGTTAGCC GTGCTAACCG TATAGCTTGC  
2651

TCCGTCGCCT CTGAGGGCAA CCACCTGCGC AGCAGGTGGG CGGCAGCCCCG CGCGCAAGCG  
2711

20 CCTACCGGGT TTGGGCACAG CCCATAAATC AACGCCTCCG GTGTTGAAGC GATCGTGTGT  
2771

25 CACGATTGCT ATGCTTGCTA CCCCTTCAGG GTTTTCTGAT ACACAAATCA AGTTTTTTTCG  
2831

TATACGCTAA TGCCATGAGT GAGCATCTAC TGCACGGCAA GCCCGTCACC AACGAGCAGA  
2891

30 TTCAGGCATG GGCAGACGAG GCCGAGGCCG GATACGACCT GCCCAAATC CCCAAGCCAC  
2951

GGCGCGGACG CCCGCCCCGTA GGAGACGGTC CGGGCACCGT CGTACCCGTG CGTCTCGACG  
3011

35 CGGCCACCGT TGCCGCTCTC ACAGAACGAG CAACAGCCGA GGGCATCACG AACCGTTTACG  
3071

40 ACGCGATCCG AGCCGCAGTC CACGAGTGA CACGGGTTGC CTGACCTCCA CGACTCAGCA  
3131

CGCAAGCACT ACCAACGAGA CCGGCTCGAC GACACGGCCG TGCTCTACGC GGCCACCCAC  
3191

45 GTTCTCAACT CCCGGCCACT CGACGACGAA GACGACCCGC GCCGCTGGCT CATGATCGGA  
3251

ACCGACCCAG CAGGCCGCCT ACTCGAACTC GTCGCACTGA TCTACGACGA CGGCTACGAA  
3311

50 CTGATCATCC ACGCAATGAA AGCCCGCACC CAATACCTCG ACCAGCTCTA ACCAAGAAAG  
3371

GAACCTGATG AGCGACCAGC TAGACAGCGA CCGCAACTAC GACCCGATGA TCTTCGACGT  
3431

5 GATGCGCGAG ACCGCGAACC GCGTCGTCGC CACGTACGTT GCATGGGAAG ATGAAGCCGC  
3491

TGATCCCCGC GAGGCTGCGC ACTGGCAGGC CGAGCGATTG CGCACCCGGC ACGAGGTGCG  
3551

10 CGCC  
3555

(2) INFORMATIONS POUR SEQ ID N° 2 :

(i) CARACTERISTIQUES DE LA SEQUENCE :

15 (A) LONGUEUR : 303 acides aminés

(B) TYPE : Acide aminé

(D) TOPOLOGIE : Linéaire

(ii) TYPE DE MOLECULE : Protéine

(xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 2 :

20

Met Asp Ser Phe Glu Thr Leu Phe Pro Glu Ser Trp Leu Pro Arg Lys  
1 5 10 15

25

Pro Leu Ala Ser Ala Glu Lys Ser Gly Ala Tyr Arg His Val Thr Arg  
20 25 30

Gln Arg Ala Leu Glu Leu Pro Tyr Ile Glu Ala Asn Pro Leu Val Met  
35 40 45

30

Gln Ser Leu Val Ile Thr Asp Arg Asp Ala Ser Asp Ala Asp Trp Ala  
50 55 60

35

Ala Asp Leu Ala Gly Leu Pro Ser Pro Ser Tyr Val Ser Met Asn Arg  
65 70 75 80

Val Thr Thr Thr Gly His Ile Val Tyr Ala Leu Lys Asn Pro Val Cys  
85 90 95

40

Leu Thr Asp Ala Ala Arg Arg Arg Pro Ile Asn Leu Leu Ala Arg Val  
100 105 110

Glu Gln Gly Leu Cys Asp Val Leu Gly Gly Asp Ala Ser Tyr Gly His  
115 120 125

45

Arg Ile Thr Lys Asn Pro Leu Ser Thr Ala His Ala Thr Leu Trp Gly  
130 135 140

Pro Ala Asp Ala Leu Tyr Glu Leu Arg Ala Leu Ala His Thr Leu Asp  
145 150 155 160

50

Glu Ile His Ala Leu Pro Glu Ala Gly Asn Pro Arg Arg Asn Val Thr  
 165 170 175  
 5 Arg Ser Thr Val Gly Arg Asn Val Thr Leu Phe Asp Thr Thr Arg Met  
 180 185 190  
 Trp Ala Tyr Arg Ala Val Arg His Ser Trp Gly Gly Pro Val Ala Glu  
 195 200 205  
 10 Trp Glu His Thr Val Phe Glu His Ile His Leu Leu Asn Glu Thr Ile  
 210 215 220  
 Ile Ala Asp Glu Phe Ala Thr Gly Pro Leu Gly Leu Asn Glu Leu Lys  
 225 230 235 240  
 15 His Leu Ser Arg Ser Ile Ser Arg Trp Val Trp Arg Asn Phe Thr Pro  
 245 250 255  
 20 Glu Thr Phe Arg Ala Arg Gln Lys Ala Ile Ser Leu Arg Gly Ala Ser  
 260 265 270  
 Lys Gly Gly Lys Glu Gly Gly His Lys Gly Gly Ile Ala Ser Gly Ala  
 275 280 285  
 25 Ser Arg Arg Ala His Thr Arg Gln Gln Phe Leu Glu Gly Leu Ser  
 290 295 300

## (2) INFORMATIONS POUR SEQ ID N° 3 :

## (i) CARACTERISTIQUES DE LA SEQUENCE :

30 (A) LONGUEUR : 85 acides aminés  
 (B) TYPE : Acide aminé  
 (D) TOPOLOGIE : Linéaire

## (ii) TYPE DE MOLECULE : Protéine

## (xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 3 :

35 Met Thr Thr Arg Glu Arg Leu Pro Arg Asn Gly Tyr Ser Ile Ala Ala  
 1 5 10 15  
 40 Ala Ala Lys Lys Leu Gly Val Ser Glu Ser Thr Val Lys Arg Trp Thr  
 20 25 30  
 Ser Glu Pro Arg Glu Glu Phe Val Ala Arg Val Ala Ala Arg His Ala  
 35 40 45  
 45 Arg Ile Arg Glu Leu Arg Ser Glu Gly Gln Ser Met Arg Ala Ile Ala  
 50 55 60

Ala Glu Val Gly Val Ser Val Gly Thr Val His Tyr Ala Leu Asn Lys  
 65 70 75 80

5 Asn Arg Thr Asp Ala  
 85

(2) INFORMATIONS POUR SEQ ID N° 4 :

(i) CARACTERISTIQUES DE LA SEQUENCE :

- 10 (A) LONGUEUR : 59 paires de bases  
 (B) TYPE : Acide nucléique  
 (C) NOMBRE DE BRINS : Un  
 (D) TOPOLOGIE : Linéaire

(ii) TYPE DE MOLECULE : ADN (synthétique)

15 (xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 4 :

AATTCAAGCT TGTCGACGTT AACCTGCAGG CATGCGGATC CGGTACCGAT ATCAGATCT  
 59

(2) INFORMATIONS POUR SEQ ID N° 5 :

20 (i) CARACTERISTIQUES DE LA SEQUENCE :

- (A) LONGUEUR : 59 paires de bases  
 (B) TYPE : Acide nucléique  
 (C) NOMBRE DE BRINS : Un  
 (D) TOPOLOGIE : Linéaire

25 (ii) TYPE DE MOLECULE : ADN (synthétique)

(xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 5 :

CCGAAGATCT GATATCGGTA CCGGATCCGC ATGCCTGCAG GTTAACGTCG ACAAGCTTG  
 59

30

(2) INFORMATIONS POUR SEQ ID N° 6 :

(i) CARACTERISTIQUES DE LA SEQUENCE :

- 35 (A) LONGUEUR : 24 paires de bases  
 (B) TYPE : Acide nucléique  
 (C) NOMBRE DE BRINS : Un  
 (D) TOPOLOGIE : Linéaire

(ii) TYPE DE MOLECULE : ADN (synthétique)

## (xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 6 :

GTACCGGCCG CTGCGGCCAA GCTT  
24

5

## (2) INFORMATIONS POUR SEQ ID N° 7 :

## (i) CARACTERISTIQUES DE LA SEQUENCE :

(A) LONGUEUR : 24 paires de bases

(B) TYPE : Acide nucléique

10

(C) NOMBRE DE BRINS : Un

(D) TOPOLOGIE : Linéaire

(ii) TYPE DE MOLECULE : ADN (synthétique)

## (xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 7 :

15. GATCAAGCTT GGCCGCAGCG GCCG  
24

## (2) INFORMATIONS POUR SEQ ID N° 8 :

## (i) CARACTERISTIQUES DE LA SEQUENCE :

20

(A) LONGUEUR : 35 paires de bases

(B) TYPE : Acide nucléique

(C) NOMBRE DE BRINS : Un

(D) TOPOLOGIE : Linéaire

(ii) TYPE DE MOLECULE : ADN (synthétique)

25

## (xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 8 :

AAACTGCAGC TGCTGGCTTG CGCCCGATGC TAGTC  
35

## 30 (2) INFORMATIONS POUR SEQ ID N° 9 :

## (i) CARACTERISTIQUES DE LA SEQUENCE :

(A) LONGUEUR : 76 paires de bases

(B) TYPE : Acide nucléique

(C) NOMBRE DE BRINS : Un

35

(D) TOPOLOGIE : Linéaire

(ii) TYPE DE MOLECULE : ADN (synthétique).

## (xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 9 :

AAACTGCAGC AGCTGGGCAG GCCGCTGGAC GGCCTGCCCT CGAGCTCGTC TAGAATGTGC  
60  
5 TGCCGATCCT GGTTC  
76

## (2) INFORMATIONS POUR SEQ ID N° 10 :

- 10 (i) CARACTERISTIQUES DE LA SEQUENCE :
- (A) LONGUEUR : 32 paires de bases
  - (B) TYPE : Acide nucléique
  - (C) NOMBRE DE BRINS : Un
  - (D) TOPOLOGIE : Linéaire
- 15 (ii) TYPE DE MOLECULE : ADN (synthétique)
- (xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 10 :

CTAGTCTAGA CACCGATGAG GAAACCCGAT GA  
32  
20

## (2) INFORMATIONS POUR SEQ ID N° 11 :

- (i) CARACTERISTIQUES DE LA SEQUENCE :
- (A) LONGUEUR : 36 paires de bases
  - (B) TYPE : Acide nucléique
  - 25 (C) NOMBRE DE BRINS : Un
  - (D) TOPOLOGIE : Linéaire
- (ii) TYPE DE MOLECULE : ADN (synthétique)
- (xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 11 :

30 CCCAAGCTTC TCGAGTCAGT GGTCGCTGGG CGCGCG  
36

## (2) INFORMATIONS POUR SEQ ID n° 12:

- 35 (i) CARACTERISTIQUES DE LA SEQUENCE :
- (A) LONGUEUR : 24 paires de bases
  - (B) TYPE: Acide nucléique
  - (C) NOMBRE DE BRINS : Un
  - (D) TOPOLOGIE : Linéaire
- (ii) TYPE DE MOLECULE: ADN (synthétique)



(xi) DESCRIPTION DE LA SEQUENCE : SEQ ID n° 12:

GGAGATCTAGATCGATATCTCGAG

24

5

(2) INFORMATIONS POUR SEQ ID n° 13:

(i) CARACTERISTIQUES DE LA SEQUENCE :

(A) LONGUEUR : 30 paires de bases

(B) TYPE: Acide nucléique

10

(C) NOMBRE DE BRINS : Un

(D) TOPOLOGIE : Linéaire

(ii) TYPE DE MOLECULE : ADN (synthétique)

15

(xi) DESCRIPTION DE LA SEQUENCE SEQ ID n° 13:

GATCCTCGAGATATCGATCTAGATCTCCGC

30

### REVENDICATIONS

1. Polynucléotide caractérisé en ce qu'il comprend une séquence capable de s'hybrider sélectivement avec
- 5 (a) la séquence SEQ ID n° 1 ou sa séquence complémentaire ;  
(b) une séquence provenant du plasmide de 3,6 kb de *Propionibacterium freudenreichii* CBS 101022 ;  
(c) une séquence provenant du plasmide de 3,6 kb de *Propionibacterium freudenreichii* CBS 101023 ; ou  
10 (d) une séquence qui code un polypeptide qui comprend SEQ ID n° 2 ou 3, une séquence d'acides aminés sensiblement homologue de celle-ci ou un fragment de l'une ou l'autre séquence.
2. Polynucléotide caractérisé en ce qu'il s'agit d'un plasmide à réplication autonome qui peut demeurer extrachromosomique à l'intérieur d'une  
15 cellule hôte, qui est issu d'un plasmide endogène de *Propionibacterium* et qui, quand il comprend un gène hétérologue, est capable d'exprimer ce gène à l'intérieur de la cellule hôte.
3. Polynucléotide selon la revendication 1, caractérisé en ce qu'il se réplique de manière autonome dans une cellule hôte.
- 20 4. Polynucléotide selon la revendication 3, caractérisé en ce que la cellule hôte est une cellule de *Propionibacterium*.
5. Polynucléotide selon la revendication 4, caractérisé en ce que la cellule de *Propionibacterium* est une cellule de *Propionibacterium freudenreichii*.
- 25 6. Polynucléotide selon l'une quelconque des revendications précédentes, caractérisé en ce qu'il est capable de s'hybrider sélectivement avec une ou plusieurs séquence de SEQ ID n° 1 qui est ou sont nécessaires pour la réplication autonome dans une cellule de *Propionibacterium*.
7. Polynucléotide selon la revendication 1, caractérisé en ce qu'il comprend le fragment de 1,7 kb de SEQ ID n° 1 délimité par les sites de  
30 restriction *SalI* et *AlwNI* ou les nucléotides 1 à 1750 de SEQ ID n° 1.
8. Vecteur caractérisé en ce qu'il comprend un polynucléotide selon l'une quelconque des revendications précédentes.
9. Vecteur selon la revendication 8, caractérisé en ce qu'il s'agit d'un plasmide.
- 35 10. Vecteur selon l'une quelconque des revendications 8 et 9, caractérisé en ce qu'il comprend en outre un marqueur sélectionnable.

11. Vecteur selon l'une quelconque des revendications 8 à 10, caractérisé en ce qu'il se réplique de manière autonome dans *E. coli*.
12. Vecteur selon l'une quelconque des revendications 8 à 11, caractérisé en ce qu'il s'agit d'un vecteur d'expression.
- 5 13. Vecteur selon la revendication 12, caractérisé en ce qu'il comprend un gène endogène de *Propionibacterium* ou un gène hétérologue lié de manière active à une séquence de contrôle capable d'assurer l'expression du gène.
14. Vecteur selon la revendication 13, caractérisé en ce que le gène est le gène *cobA*.
- 10 15. Vecteur selon la revendication 13, caractérisé en ce que le gène hétérologue code un polypeptide à caractère thérapeutique chez un être humain ou un animal.
16. Polypeptide caractérisé en ce qu'il comprend la séquence SEQ ID n° 2 ou 3 ou une séquence sensiblement homologue à celle-ci, ou un fragment de
- 15 l'une ou l'autre desdites séquences, ou en ce qu'il est codé par un polynucléotide selon l'une quelconque des revendications 1 à 7.
17. Cellule hôte, caractérisée en ce qu'elle comprend un polynucléotide selon l'une quelconque des revendications 1 à 7 ou un vecteur selon l'une quelconque des revendications 8 à 15 ou en ce qu'elle peut être
- 20 transformée ou transfectée avec un vecteur selon l'une quelconque des revendications 13 à 15.
18. Cellule hôte selon la revendication 17, caractérisée en ce qu'il s'agit d'une bactérie.
19. Cellule hôte selon la revendication 18, caractérisée en ce qu'il
- 25 s'agit d'une bactérie *Propionibacterium* ou *E. coli*.
20. Procédé de production d'une cellule hôte selon l'une quelconque des revendications 17 à 19, caractérisé en ce qu'il comprend la transformation ou la transfection d'une cellule hôte avec un polynucléotide selon l'une quelconque des revendications 1 à 7 ou avec un vecteur selon l'une quelconque des
- 30 revendications 8 à 15.
21. Procédé de préparation d'un polypeptide ou d'un autre composé, caractérisé en ce qu'il comprend la culture ou la fermentation d'une cellule hôte selon l'une quelconque des revendications 17 à 19 dans des conditions permettant l'expression ou la production du polypeptide ou du composé.

22. Procédé selon la revendication 21, caractérisé en ce qu'il s'agit d'un procédé de fermentation dans lequel la cellule hôte est cultivée dans des conditions aérobies ou anaérobies.

23. Procédé selon l'une quelconque des revendications 21 et 22, caractérisé en ce que le polypeptide exprimé ou le composé produit est récupéré à partir de la cellule hôte.

24. Procédé selon la revendication 23, caractérisé en ce que le polypeptide est une protéase, une amylase, une lipase ou une peptidase, ou en ce que le composé est la vitamine B12.

25. Procédé selon l'une quelconque des revendications 21 à 24, caractérisé en ce que le polypeptide est sécrété par la cellule hôte.

26. Procédé selon l'une quelconque des revendications 21 à 24, caractérisé en ce que le polypeptide est exprimé sur la surface de la cellule hôte et en ce qu'il s'agit d'un antigène ou d'un immunogène.

27. Polypeptide ou composé, caractérisé en ce qu'il est préparé par un procédé selon l'une quelconque des revendications 21 à 26.

28. Procédé de production de la vitamine B12 (cobalamine), caractérisé en ce qu'il comprend la culture d'une cellule hôte selon l'une quelconque des revendications 17 à 19 dans des conditions dans lesquelles la vitamine B12 est produite et, si nécessaire, l'isolement de la vitamine B12.

29. Vitamine B12, caractérisée en ce qu'elle est produite par le procédé selon la revendication 28.

30. Polypeptide selon la revendication 27, caractérisé en ce qu'il est destiné à être utilisé dans une méthode de traitement du corps humain ou animal par thérapie.

31. Cellule hôte selon l'une quelconque des revendications 17 à 19, caractérisée en ce qu'elle est destinée à être utilisée dans une méthode de traitement du corps humain ou animal par thérapie ou dans un aliment pour animaux.

32. Utilisation d'une cellule hôte selon l'une quelconque des revendications 17 à 19 ou d'un polypeptide ou composé selon la revendication 27 pour la fabrication du fromage.

33. Utilisation d'une cellule hôte selon l'une quelconque des revendications 17 à 19 ou d'un polypeptide ou composé selon la revendication 27 dans la fabrication d'un aliment ou d'un aliment pour animaux.

34. Aliment, caractérisé en ce qu'il comprend un polypeptide ou composé selon la revendication 27 ou une cellule hôte selon l'une quelconque des revendications 17 à 19.

5 35. Aliment selon la revendication 34, caractérisé en ce qu'il est destiné à la consommation par l'homme (par exemple fromage, saucisse) ou par un animal.

36. Procédé de fabrication du fromage ou d'autres produits laitiers fermentés, caractérisé en ce qu'il comprend l'utilisation d'une cellule hôte selon l'une quelconque des revendications 17 à 19.

10 37. Procédé selon la revendication 36, caractérisé en ce que la cellule hôte est utilisée à la place de ou en plus de bactéries lactiques.

38. Procédé selon l'une quelconque des revendications 36 et 37, caractérisé en ce que la cellule hôte est une cellule de *Propionibacterium*.

Figure 1

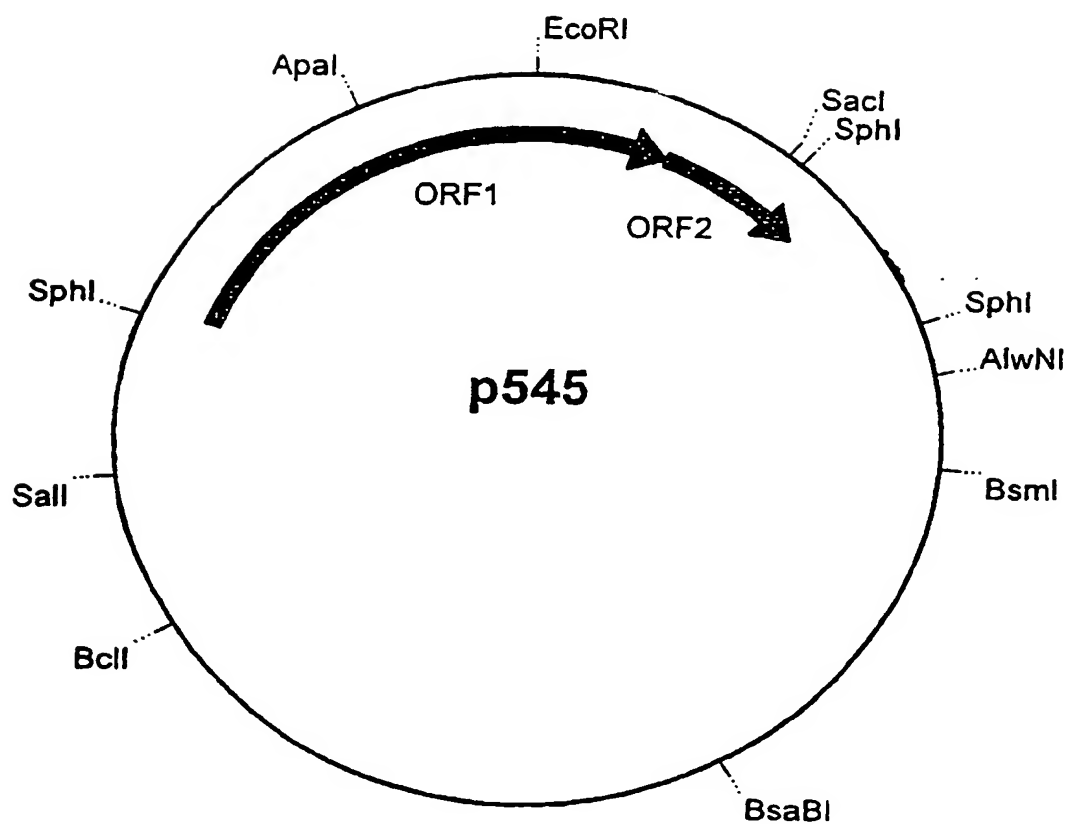


Figure 2a

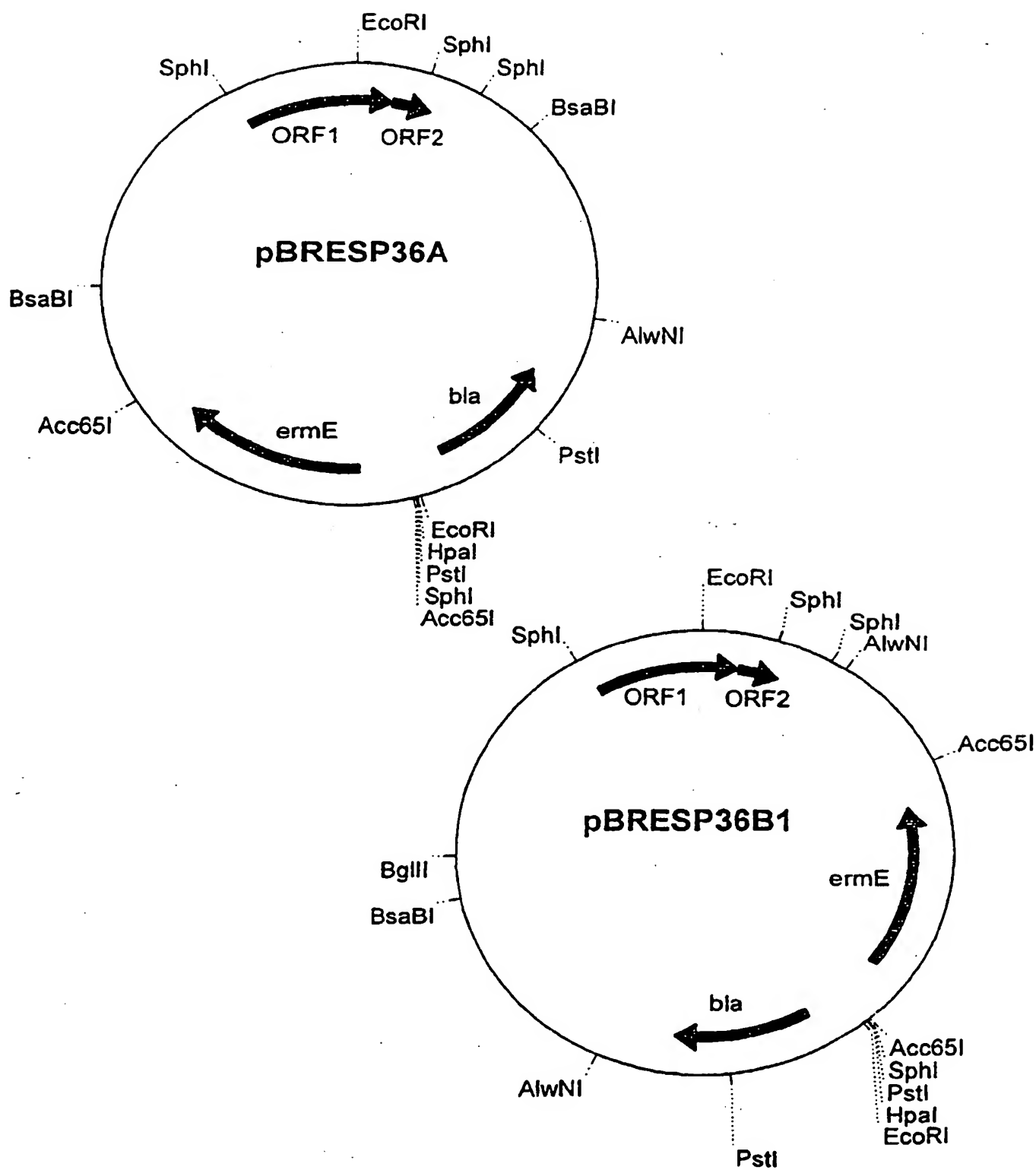


Figure 2b

